



**Liliana da Silva
Ferreira**

**Contributos para a Caracterização das Vogais
Nasais do Português**



**Liliana da Silva
Ferreira**

**Contributos para a Caracterização das Vogais
Nasais do Português**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Electrónica e Telecomunicações, realizada sob a orientação científica do Doutor António Joaquim da Silva Teixeira, Professor do Departamento de Engenharia Electrónica e Telecomunicações da Universidade de Aveiro.

O júri

Presidente

Prof. Doutor Francisco António Cardoso Vaz
Professor Catedrático da Universidade de Aveiro

Vogais

Prof. Doutor Diamantino António Caseiro
Professor Auxiliar do Instituto Superior Técnico da Universidade Técnica de Lisboa

Prof. Doutor António Joaquim da Silva Teixeira
Professor Auxiliar da Universidade de Aveiro

Agradecimentos / acknowledgments

Durante o desenvolvimento deste trabalho recebi o apoio e a colaboração de diversas pessoas, às quais gostaria de transmitir o meu agradecimento geral. Agradeço, em particular, ao Prof. António Teixeira, pela orientação e pela disponibilidade constante ao longo do trabalho.

Resumo

As vogais nasais do Português Europeu são sons complexos com uma natureza dinâmica que apresentam várias dificuldades à sua caracterização completa.

Em estudos anteriores, visando a caracterização destes sons nasais, foram recolhidos dois *corpora* contendo, para além do sinal de voz, informação sobre os movimentos orais e do véu palatino (recolhidos através da tecnologia *Electromagnetic Articulography*) e informação sobre o comportamento das cordas vocais (sinal proveniente do *Electroglottograph*), respectivamente. Neste último *corpus* foram ainda contempladas as variações regionais e contextuais das vogais nasais.

Este trabalho tem como objectivo estender as análises efectuadas anteriormente a todo o material de voz existente, melhorar a obtenção de parâmetros relacionados com a fonte e caracterizar as várias fases de uma vogal nasal no maior número de contextos possível.

As análises desenvolvidas serão úteis na produção das várias vogais nasais do Português, utilizando o sintetizador articulatório da Universidade de Aveiro, com qualidade, se possível, superior à resultante de estudos anteriores.

Abstract

The European Portuguese nasal vowels are complex sounds with a dynamic nature that present several difficulties to its complete characterization.

In previous studies, aiming the analysis of these nasal sounds, were collected two *corpora* containing, beyond the speech signal, information about the oral and velum movements (collected with *Electromagnetic Articulography*), and information about the glottal source behaviour (signal originating from the *Electroglottograph*), respectively. In this last *corpus* were contemplated the regional and contextual variations of the nasal vowels.

This study extend the previous analysis to all the existing speech material, improve the obtaining of parameters related with the glottal source and characterize the several phases of a nasal vowel in all possible number of contexts. The analysis developed will be helpful in the production of the several nasal vowels of the Portuguese, if possible, with superior quality over the resultant one from previous studies, utilizing for such the University of Aveiro articulatory synthesizer.

Índice

1	Introdução	1
1.1	Introdução	1
1.2	Objectivos	2
1.3	Resultados Publicados	3
1.4	Estrutura da Dissertação	3
2	Vogais Nasais	5
2.1	A Nasalidade	5
2.1.1	Processo de Produção de Fala	5
2.1.2	Anatomia e Fisiologia das Cavidades Nasais	8
2.2	O Sistema Vocálico Nasal do Português	10
2.3	Propriedades Acústicas das Vogais Nasais	10
2.3.1	Percepção de Vogais Nasais	11
2.4	Estudos de Vogais Nasais	12
2.4.1	Técnicas Utilizadas	13
2.4.2	Estudos das Vogais Nasais Portuguesas	14
3	<i>Corpora</i>	19
3.1	<i>Corpus</i> EMMA	19
3.1.1	Constituição do <i>Corpus</i>	20
3.1.2	Recolha e Informante	21
3.2	<i>Corpus</i> Regional	23
3.2.1	Constituição do <i>Corpus</i>	23
3.2.2	Locais de Recolha e Características dos Informantes	24
3.2.3	Software Utilizado	26
3.2.4	Segmentação e etiquetagem	26
4	Características da Fonte	29
4.1	Introdução	30
4.2	Método	32
4.2.1	Frequência Fundamental	32
4.2.2	Quociente de Abertura	33
4.2.3	<i>Jitter</i>	34

4.3	Resultados	34
4.3.1	Frequência Fundamental	34
4.3.2	Quociente de Abertura	37
4.3.3	<i>Jitter</i>	38
4.4	Análise após normalização	43
4.4.1	Método	45
4.4.2	Resultados	45
4.5	Discussão	47
4.6	Resumo	48
5	Caracterização Temporal	49
5.1	Duração Total	50
5.1.1	Introdução	50
5.1.2	<i>Corpus</i> Regional	50
5.1.3	<i>Corpus</i> EMMA	55
5.1.4	Discussão	56
5.2	Obtenção de informação acerca das Fases de uma vogal nasal	57
5.2.1	Utilizando anotações dos sensores EMMA	58
5.2.2	Com base apenas no sinal acústico: estudo exploratório	61
5.3	Resumo	65
6	Características Relacionadas com o Tracto	67
6.1	Introdução	67
6.2	Análise da Fase Inicial de uma Vogal Nasal no Contexto $C\tilde{V}C$	67
6.2.1	Método	68
6.2.2	Resultados	69
6.2.3	Discussão	75
6.3	Análise da Fase Final	77
6.3.1	Introdução	77
6.3.2	Objectivos	80
6.3.3	Método	80
6.3.4	Resultados	84
6.3.5	Discussão	87
6.4	Resumo	88
7	Conclusões	89
7.1	Resumo do Trabalho	89
7.2	Resultados Importantes	90
7.3	Trabalho Futuro	90
A	Constituição do <i>corpus</i> Regional	97
B	Características dos Informantes	103

C	Estímulos Visuais	105
D	Resultados Características da Fonte sem médias	109
D.1	Frequência Fundamental	109
D.2	Quociente de Abertura	109
E	Resultados da Caracterização das Fases das Vogais Nasais com Base no Sinal Acústico	119
F	Resultados da Análise da Fase Inicial de uma Vogal Nasal	125

Lista de Figuras

2.1	Aparelho fonador.	6
2.2	Corte sagital da cavidade nasal.	8
2.3	Seios Paranasais.	9
3.1	Diagrama mostrando a posição dos vários sensores utilizados.	21
3.2	Exemplo retirado do <i>corpus</i> EMMA.	22
3.3	Exemplo mostrando o EGG (na esquerda) e o sinal de voz (traçado superior) e o sinal proveniente do EGG (em baixo).	24
3.4	Mapa de Portugal com indicação das zonas onde pertence cada um dos informantes.	25
3.5	Figuras utilizadas na gravação do <i>corpus</i> . As imagens correspondem às palavras <i>ponte</i> , <i>pintor</i> e <i>bandeira</i> , respectivamente.	26
3.6	Exemplo de uma anotação do <i>corpus</i> . De cima para baixo, pode ver-se o sinal de voz, o sinal do EGG, anotação ao nível das palavras e anotação fonética. Nesta última pode ver-se os pontos de início da vogal nasal e da consoante que a segue ([t]) assim como a indicação de qual o segmento que antecede a vogal nasal (neste caso o [p]).	27
3.7	Distribuição das amostras anotadas por vogal e por região.	28
4.1	Exemplo de cálculo de F0 com base no sinal EGG. Através do sinal do EGG, apresentado no topo, é primeiro calculado o início de cada período glotal (<i>tx</i> no sistema SFS) e depois calculada a frequência fundamental (<i>fx</i> no sistema SFS). No final é visível a frequência fundamental obtida através do sinal de voz.	33
4.2	Intervalo de Confiança para os valores de F0 (em Hz) em função da vogal, da nasalidade (nasais em baixo), do acento (acentuadas nos dois gráficos da direita).	35
4.3	Intervalo de Confiança para os valores de F0 (Hz) das vogais nasais (acentuadas ou não) em função da vogal e do falante.	37
4.4	Diagrama de extremos e quartis (<i>boxplot</i>) dos valores do quociente de abertura em função da vogal e nasalidade.	38
4.5	<i>Boxplots</i> do quociente de abertura para cada falante.	38
4.6	<i>Boxplots</i> do valor do Jitter Factor na posição inicial da vogal. São apresentados valores para cada vogal separados pela nasalidade.	41

4.7	<i>Boxplots</i> do valor do Jitter Factor na posição medial da vogal. São apresentados valores para cada vogal separados pela nasalidade.	41
4.8	<i>Boxplots</i> do valor do Jitter Factor na posição final da vogal. São apresentados valores para cada vogal separados pela nasalidade.	42
4.9	<i>Boxplots</i> do valor do Jitter Factor usando a totalidade da duração da vogal. São apresentados valores para cada vogal separados pela nasalidade.	42
4.10	Numerador do cálculo de JF para orais.	44
4.11	Numerador do cálculo de JF para nasais.	44
4.12	Diagrama de extremos e quartis dos valores de F0MED em função do falante.	46
4.13	Diagrama de extremos e quartis dos valores de F0MED em função das vogais nasais após divisão pela mediana de F0MED.	46
4.14	Diagrama de extremos e quartis dos valores de OQMean em função das vogais nasais após divisão pela mediana de OQMean.	47
5.1	Intervalo de confiança a 95% dos valores médios da duração (ms) das vogais nasais do <i>corpus</i> Regional.	51
5.2	Intervalo de Confiança a 95% e valor médio da duração das vogais orais e nasais (ms). Os dois gráficos de cima referem-se a vogais orais (nasal=0), os dois gráficos da esquerda referem-se ao caso não acentuado (accent=0).	53
5.3	Intervalo de Confiança das durações (ms) para as vogais nasais em posição não acentuada (em cima) e acentuada (em baixo) apresentando separadamente os valores obtidos para cada região.	54
5.4	Intervalo de confiança a 95% para os valores médios das durações das vogais nasais do <i>corpus</i> EMMA (s).	56
5.5	Intervalo de confiança a 95% para os valores médios das durações das vogais nasais dos informantes masculinos, da região da Beira Litoral, do <i>corpus</i> Regional (ms).	56
5.6	Intervalo de confiança a 95% para os valores médios das durações das vogais nasais dos informantes masculinos, da região do Minho, do <i>corpus</i> Regional (ms).	57
5.7	Eventos do velo e orais para uma vogal nasal entre consoantes orais.	58
5.8	<i>Scatter Plot</i> entre o evento de abertura do velo (s) e o primeiro segmento automático (s).	62
5.9	<i>Scatter Plot</i> entre o evento de fecho dos articuladores orais (s) e o segundo segmento automático (s).	63
5.10	<i>Scatter Plot</i> entre o evento de abertura do velo (s) e o primeiro segmento automático (em 4 segmentos) (s).	64
5.11	<i>Scatter Plot</i> entre o evento de fecho dos articuladores orais e o quarto segmento automático (em 4 segmentos).	65
6.1	Representação da distribuição das duas primeiras formantes para a totalidade dos falantes, separando os valores pelo sexo (Homens do lado esquerdo) e existência ou não de nasalidade (nasais no topo) (Hz).	70

6.2	Boxplots de F1 (separação por sexo e nasalidade) (Hz).	71
6.3	Boxplots de F2 (separação por sexo e nasalidade) (Hz).	72
6.4	Boxplots da diferença entre F2 e F1 (separação por sexo e nasalidade) (Hz).	73
6.5	Triângulos acústicos para as vogais nasais das seis regiões contempladas no <i>corpus</i> (Hz).	74
6.6	Triângulos acústicos para as vogais orais das seis regiões contempladas no <i>corpus</i> (Hz).	75
C.1	Imagem correspondente à palavra <i>Candeeiro</i>	105
C.2	Imagem correspondente à palavra <i>Tambor</i>	105
C.3	Imagem correspondente à palavra <i>Cambalhota</i>	106
C.4	Imagem correspondente à palavra <i>Estendal</i>	106
C.5	Imagem correspondente à palavra <i>Pombo</i>	106
C.6	Imagem correspondente à palavra <i>Comboio</i>	107
C.7	Imagem correspondente à palavra <i>Pandeireta</i>	107
C.8	Imagem correspondente à palavra <i>Pinguim</i>	107
D.1	Diagrama de extremos e quartis dos valores de F0MED em função do falante após subtração da mediana de F0MED.	109
D.2	Diagrama de extremos e quartis dos valores de F0MED em função da vogal após subtração da mediana de F0MED.	110
D.3	Diagrama de extremos e quartis dos valores de F0MED em função das vogais nasais após subtração da mediana de F0MED.	110
D.4	Diagrama de extremos e quartis dos valores de F0MED em função da vogal após subtração da mediana de F0MED e divisão por F0MED.	111
D.5	Diagrama de extremos e quartis dos valores de F0MED em função das vogais nasais após subtração da mediana de F0MED e divisão por F0MED.	111
D.6	Diagrama de extremos e quartis dos valores de F0START em função do falante.	112
D.7	Diagrama de extremos e quartis dos valores de F0START em função da vogal após subtração da mediana de F0START.	112
D.8	Diagrama de extremos e quartis dos valores de F0START em função das vogais nasais após subtração da mediana de F0START.	113
D.9	Diagrama de extremos e quartis dos valores de F0END em função do falante.	113
D.10	Diagrama de extremos e quartis dos valores de F0END em função da vogal após subtração da mediana de F0END.	114
D.11	Diagrama de extremos e quartis dos valores de F0END em função das vogais nasais após subtração da mediana de F0END.	114
D.12	Diagrama de extremos e quartis dos valores de OQMean em função da vogal após subtração da mediana de OQMEan.	115
D.13	Diagrama de extremos e quartis dos valores de OQMean em função das vogais nasais após subtração da mediana de OQMEan.	116

D.14	Diagrama de extremos e quartis dos valores de OQMean em função das vogais nasais após subtração da mediana de OQMean e divisão por OQMean. . . .	117
E.1	<i>Scatter Plot</i> entre o evento médio do velo e o segundo segmento automático (em três).	119
E.2	<i>Scatter Plot</i> entre o evento de fecho do velo e o terceiro segmento automático (em três).	120
E.3	<i>Scatter Plot</i> entre o evento de abertura dos articuladores orais e o segundo segmento automático (em três).	121
E.4	<i>Scatter Plot</i> entre o evento de fecho do velo e o quarto segmento automático (em 4 segmentos).	121
E.5	<i>Scatter Plot</i> entre o evento de fecho dos articuladores orais e o primeiro segmento automático (em 4 segmentos).	122
E.6	<i>Scatter Plot</i> entre o evento médio do velo e o segundo segmento automático (em 4 segmentos).	123
E.7	<i>Scatter Plot</i> entre o evento medial do velo e o quarto segmento automático (em 4 segmentos).	124
F.1	Valores de F1 e F2 (à direita) para as vogais [6̃][6] e [a].	125
F.2	Valores de F1 e F2 (à direita) para as vogais [ẽ][e] e [E].	126
F.3	Valores de F1 e F2 (à direita) para as vogais [ĩ] e [i].	126
F.4	Valores de F1 e F2 (à direita) para as vogais [õ][o] e [O].	127
F.5	Valores de F1 e F2 (à direita) para as vogais [ũ] e [u].	127
F.6	Representação da distribuição das duas primeiras formantes para os falantes do MINHO.	128
F.7	Boxplots de F1 (separação por sexo e nasalidade) para os falantes do MINHO.	128
F.8	Boxplots de F2 (separação por sexo e nasalidade) para os falantes do MINHO.	129
F.9	Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes do MINHO.	129
F.10	Representação da distribuição das duas primeiras formantes para os falantes de TRÁS-OS-MONTES.	130
F.11	Boxplots de F1 (separação por sexo e nasalidade) para os falantes de TRAS-OS-MONTES.	130
F.12	Boxplots de F2 (separação por sexo e nasalidade) para os falantes de TRAS-OS-MONTES.	131
F.13	Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes de TRAS-OS-MONTES.	131
F.14	Representação da distribuição das duas primeiras formantes para os falantes da BEIRA LITORAL.	132
F.15	Boxplots de F1 (separação por sexo e nasalidade) para os falantes da BEIRA LITORAL.	133
F.16	Boxplots de F2 (separação por sexo e nasalidade) para os falantes da BEIRA LITORAL.	133

F.17 Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes da BEIRA LITORAL.	134
F.18 Representação da distribuição das duas primeiras formantes para os falantes da BEIRA INTERIOR.	134
F.19 Boxplots de F1 (separação por sexo e nasalidade) para os falantes da BEIRA INTERIOR.	135
F.20 Boxplots de F2 (separação por sexo e nasalidade) para os falantes da BEIRA INTERIOR.	135
F.21 Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes da BEIRA INTERIOR.	136
F.22 Representação da distribuição das duas primeiras formantes para os falantes do ALENTEJO.	136
F.23 Boxplots de F1(separação por sexo e nasalidade) para os falantes do ALENTEJO.	137
F.24 Boxplots de F2 (separação por sexo e nasalidade) para os falantes do ALENTEJO.	137
F.25 Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes do ALEN-TEJO.	138
F.26 Representação da distribuição das duas primeiras formantes para os falantes do MINHO.	138
F.27 Boxplots de F1 (separação por sexo e nasalidade) para os falantes do ALGARVE.	139
F.28 Boxplots de F2 (separação por sexo e nasalidade) para os falantes do ALGARVE.	139
F.29 Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes do ALGARVE.	140

Lista de Tabelas

3.1	Os vários contextos contemplados no <i>corpus</i> regional.	23
3.2	Excerto das palavras seleccionadas para o contexto vogal nasal entre oclusivas não vozeadas.	24
4.1	Valores médios de F0 (Hz) medidos no início, meio e fim da vogal para cada um dos 25 informantes.	36
4.2	Diferença entre os valores inicial e final de F0 (Hz).	36
4.3	Valores Médios de OQ para as vogais orais e por falante. Na primeira coluna apresenta-se a sigla pela qual é identificado o falante.	39
4.4	Valores Médios de OQ para vogais nasais e por falante.	40
4.5	Valores médios obtidos para o Jitter Factor. Na primeira tabela por vogal e tendo em conta a nasalidade. Na tabela intermédia médias globais para os valores medidos nas três posições e na globalidade da vogal. Na tabela de baixo à esquerda os valores para vogais orais, à direita para vogais nasais.	43
5.1	Duração Total (ms).	51
5.2	Durações médias das vogais analisadas (ms).	52
5.3	Médias de duração para as vogais nasais das 6 regiões, a negrito encontram-se os valores superiores (ms).	52
5.4	Duração Total (ms).	55
5.5	Durações Fases (ms).	59
5.6	Durações Fases (%).	60
5.7	<i>Ratio</i> r entre d0 e dc.	60
5.8	Sumário do Modelo obtido para o evento de abertura do velo e o primeiro segmento automático.	62
5.9	Sumário do Modelo obtido para o evento de fecho dos articuladores orais (s) e o segundo segmento automático (s).	63
5.10	Sumário do Modelo obtido para o evento de abertura nasal e o primeiro segmento automático (em quatro segmentos)	64
5.11	Sumário do Modelo obtido para o evento de fecho oral e o quarto segmento automático.	64

6.1	Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [6̃].	84
6.2	Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [ẽ].	85
6.3	Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [ĩ].	86
6.4	Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [õ].	86
6.5	Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [ũ].	87
6.6	Resumo quantitativo dos valores atribuídos por cada medida de distância espectral à distância entre a mesma vogal ($V_i V_i$).	87
A.1	Vogais Nasais entre oclusivas surdas.	97
A.2	Vogais Nasais entre oclusivas sonoras.	98
A.3	Vogais Nasais após oclusivas surdas e antes de oclusivas sonoras.	98
A.4	Vogais Nasais após oclusivas sonoras e antes de oclusivas surdas.	99
A.5	Vogais Nasais entre fricativas.	99
A.6	Palavras com Consoantes Nasais.	99
A.7	Vogais Nasais antes ou após uma oclusiva.	100
A.8	Vogais Nasais após consoante líquida.	100
A.9	Vogais Orais em vários contextos.	101
A.10	Vogais Nasais em início ou fim de palavra.	101
A.11	Sequências contendo vogais nasais nos limites das palavras.	102
A.12	Ditongos Nasais.	102
B.1	Informantes do <i>corpus</i> Regional.	103
E.1	Sumário do Modelo obtido para o evento médio do velo e o segundo segmento automático (em três).	119
E.2	Sumário do Modelo obtido para o evento de fecho do velo e o terceiro segmento automático (em três).	120
E.3	Sumário do Modelo obtido para o evento de abertura dos articuladores orais e o segundo segmento automático (em três).	120
E.4	Sumário do Modelo obtido para o evento de fecho do velo e o quarto segmento automático (em 4 segmentos).	122
E.5	Sumário do Modelo obtido para o evento de fecho dos articuladores orais e o primeiro segmento automático (em 4 segmentos).	122
E.6	Sumário do Modelo obtido para o evento médio do velo e o segundo segmento automático (em 4 segmentos).	123
E.7	Sumário do Modelo obtido para o evento medial do velo e o quarto segmento automático (em 4 segmentos).	123

Glossário

CLC Centro de Línguas e Culturas.

CV Coeficiente de Variação.

DAT *Digital Audio Tape*.

EGG Electroglotógrafo / Electroglotograma.

EMMA *ElectroMagnetic Midsagittal Articulography*.

EMU *EMU Speech Database System*.

F0 Frequência Fundamental.

F1 Primeira Formante.

F2 Segunda Formante.

JF *Jitter Factor*.

KTH *Kungliga Tekniska högskolan (Royal Institute of Technology)*.

LREC *International Conference on Language Resources and Evaluation*.

IEETA Instituto de Engenharia Electrónica e Telemática de Aveiro.

MFCC *Mel Frequency Cepstral Coefficients*.

MIT *Massachusetts Institute of Technology*.

PE Português Europeu.

PLP *Perceptual Linear Prediction*.

PVI *Period Variability Index*.

OQ Quociente de Abertura.

SAMPA *Speech Assessment Methods Phonetic Alphabet*.

SFS *Speech Filing System*.

SML *Speech Measurement Language*.

Capítulo 1

Introdução

1.1 Introdução

Desde 1995 que se trabalha, no Departamento de Electrónica e Telecomunicações da Universidade de Aveiro, na síntese dos sons nasais do Português. A principal motivação prende-se com a obtenção de síntese com qualidade próxima da natural. Um dos caminhos mais promissores para atingir esse objectivo é a utilização de síntese articulatória, técnica que se baseia na modelação directa do processo de produção humano. A síntese articulatória permite realizar simulações e, funcionando como um falante versátil, obter estímulos para a realização de testes perceptuais, sendo útil em estudos indispensáveis para a obtenção de novos conhecimentos sobre os mecanismos de produção e percepção.

Assim, e em vez de dispersar a atenção por todos os tipos de sons da língua portuguesa, optou-se por uma dedicação prioritária ao estudo das vogais nasais do Português Europeu (PE).

As vogais nasais são sons complexos com uma natureza dinâmica que apresentam várias dificuldades à sua caracterização completa. Em parte devido a estas dificuldades existem algumas áreas em que é necessária mais informação. Áreas como a variação regional, ou não, das características das vogais nasais do PE, o comportamento da fonte glotal na produção destes sons, ou ainda a forma como as vogais nasais são produzidas e percebidas, são áreas pouco exploradas e que carecem de informação e estudo.

Neste sentido, foram recolhidos dois novos *corpora* especialmente vocacionados para o estudo dos sons nasais do PE. Foi recolhida informação sobre a posição da língua, lábios e véu palatino, durante a produção de palavras e frases contendo nasais, utilizando um sistema de *ElectroMagnetic Midsagittal Articulography* (EMMA) (Teixeira e Vaz, 2001). Foi, desse modo, obtido um *corpus* de aproximadamente uma hora com medição simultânea dos movimentos do véu palatino e orais. No entanto, esta técnica não é viável para um número elevado de oradores o que faz com que o sinal de voz, eventualmente complementado com sinais de fácil aquisição, tais como o do Electroglotógrafo / Electroglotograma (EGG), seja a única informação passível de obtenção em estudos contemplando a variação regional.

Com este objectivo, foi adquirido, no âmbito do projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português”¹, tendo por parceiros o Centro de Línguas e Culturas (CLC) da Universidade de Aveiro e o Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA), um novo *corpus* contemplando as várias regiões de Portugal Continental (Alentejo, Algarve, Beira Interior, Beira Litoral, Minho e Trás-os-Montes) contendo, quer o sinal de voz, quer o sinal do EGG.

1.2 Objectivos

O trabalho efectuado continua a análise relativa aos sons nasais do PE efectuada no projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português”. Durante o projecto não foi possível efectuar análises a todo o material de voz recolhido, logo o presente trabalho pretende estender o estudo a todos os dados.

Para além deste alargamento, foram ainda desenvolvidas análises contemplando novas questões, como a caracterização das várias fases das vogais nasais, nos diferentes contextos em que estas podem ocorrer.

Com estas análises pretende-se o melhoramento do conhecimento sobre os sons nasais, em particular as vogais. O principal objectivo é a produção das várias vogais nasais do PE, usando o sintetizador articulatório da Universidade de Aveiro com qualidade, se possível, superior à resultante de trabalhos anteriores. No entanto, existem outras aplicações, tais como, a utilização do conhecimento adquirido para a melhoria do ensino da Língua Portuguesa e para materiais de terapia de fala, entre outros.

A primeira abordagem prende-se com o problema do comportamento da fonte glotal na produção dos sons nasais, alargando os estudos efectuados em fases anteriores do projecto, a todo o *corpus* Regional. O segundo problema tratado é relativo à caracterização temporal e foca, quer a duração total das vogais nasais, quer a duração das diferentes fases que as constituem. Esta problemática inclui também um estudo exploratório cujo objectivo passa por obter, de uma forma automática e com base apenas no sinal acústico, informação sobre as várias fases de uma vogal nasal. O último problema pretende a obtenção de informação relativa à configuração assumida pelo tracto durante as fases inicial e final de uma vogal nasal.

Todos os estudos foram desenvolvidos no contexto vogal nasal entre consoantes oclusivas (CVC) e, para tal, foram utilizados dois *corpora*: um contemplando as variações regionais e contextuais das vogais nasais do PE (*corpus* Regional) e outro recolhido através da tecnologia *Electromagnetic Articulography* (*corpus* EMMA) que permite analisar os movimentos dos articuladores na produção de nasalidade.

¹<http://www.ii.ua.pt/cidlc/gcl/fap/>

1.3 Resultados Publicados

Os resultados que foram obtidos durante a realização deste trabalho foram já assunto de algumas publicações.

Durante a colaboração no projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português” foram elaborados dois relatórios técnicos que contêm informação sobre o *corpus* Regional desenvolvido. O primeiro relatório, denominado “Novo *Corpus* Acústico para o Estudo das Vogais Nasais do Português Europeu Contemplando Variedades Regionais e Contextuais” tinha como principal objectivo apresentar um conjunto de dois DVDs que haviam sido criados. Estes DVDs contêm todos os dados sobre o *corpus* reunidos até à data.

O segundo relatório foi denominado “Estudo da Posição da Língua no Início das Vogais Nasais de Seis Regiões de Portugal pela Análise das Duas Primeiras Formantes”. Este relatório apresenta informação sobre o processo de obtenção das formantes, bem como a descrição de todo o estudo efectuado. Este estudo foi realizado contemplando, separadamente o género e a nasalidade. Foram considerados, ainda, os dados relativos ao conjunto das várias regiões e a informação obtida para cada uma destas.

Na *International Conference on Language Resources and Evaluation* (LREC) 2003, foi publicada uma comunicação que foca todo processo de criação do *corpus*, desde a recolha e cobertura regional, até à segmentação e anotação dos ficheiros de som recolhidos. Nesta publicação constam ainda algumas análises críticas dos resultados obtidos no estudo da posição da língua, referido anteriormente.

Recentemente, no XX Encontro Nacional da Associação Portuguesa de Linguística, foi efectuada uma comunicação que partilha o título desta dissertação. A comunicação contém alguns resultados preliminares sobre o trabalho aqui apresentado. Na altura da submissão apenas tinham sido obtidos os resultados referentes à caracterização temporal das vogais nasais e relativos à configuração durante as fases inicial e final destes sons.

1.4 Estrutura da Dissertação

Após uma introdução ao assunto e objectivos propostos, que se efectua neste capítulo, são apresentadas, no capítulo dois, as vogais nasais. Este segundo capítulo começa com uma breve descrição dos órgãos intervenientes no processo de produção dos sons nasais e segue com informação sobre o sistema vocálico. Na parte final do capítulo são destacados alguns dos estudos existentes sobre vogais nasais, reservando uma secção para a descrição de alguns trabalhos relevantes sobre as nasais portuguesas.

O material utilizado neste trabalho provém essencialmente de dois *corpora* reunidos com o objectivo de estudar as vogais nasais do PE: o *corpus* EMMA e o *corpus* Regional. No capítulo três descreve-se detalhadamente o conteúdo destes conjuntos de dados, focando aspectos que vão desde os métodos de recolha até à sua segmentação e processamento.

No capítulo quatro é apresentado o estudo desenvolvido com o objectivo de caracterizar a fonte glotal durante a produção de uma vogal nasal. Assim, este capítulo pretende a obtenção de informação sobre o efeito do acoplamento nasal na frequência fundamental, quociente de abertura e *jitter*. Deste modo, e após uma breve introdução ao assunto, são apresentados os métodos utilizados na análise, seguindo-se os resultados obtidos e sua discussão.

O capítulo cinco diz respeito à caracterização temporal das vogais nasais. Inicialmente é realizada uma análise da duração total destes fonemas nasais, a que se segue um estudo da duração de cada uma das suas fases. A parte final do capítulo apresenta o estudo exploratório efectuado com o objectivo de obter informação sobre as fases de uma vogal nasal, baseando a análise apenas no sinal acústico. Termina com uma breve discussão dos assuntos desenvolvidos.

As características das vogais nasais relacionadas com o tracto são analisadas no capítulo seis. O estudo destas características foi dividido em duas zonas, o que se reflecte na sua estrutura. Assim, começa por apresentar o estudo efectuado para a caracterização no início da vogal nasal. O objectivo e o método utilizados são apresentados, a que se segue uma descrição dos principais resultados obtidos e uma breve discussão destes. A caracterização da parte final das vogais nasais preenche as secções seguintes, cuja estrutura é semelhante à da secção anterior.

Este documento é finalizado no capítulo sete. O capítulo final resume o trabalho apresentado e os principais resultados obtidos. Este apresenta ainda algumas sugestões para trabalho futuro.

Capítulo 2

Vogais Nasais

A nasalidade, considerada sob o ponto de vista fonológico é uma das características que diferenciam o sistema vocálico português do sistema vocálico de muitos outros idiomas.

Em PE são referidas, em geral, cinco vogais nasais¹ ([ɐ̃], [ẽ], [ĩ], [õ], [ũ]), vários ditongos nasais (por exemplo, [ẽj̃], [õj̃], [ũj̃]) e até tritongos nasais (tal como [wɐ̃w̃]). Nos dialectos do norte do país é ainda possível encontrar vogais nasais baixas tais como [ã], [ɛ̃] e [ɔ̃] (Sampson, 1999).

Este capítulo² pretende, assim, introduzir estes sons, explorando assuntos que vão desde aspectos gerais relativos ao processo de produção, até outras pesquisas realizadas para o PE, que nos dão indicações importantes acerca das características das vogais nasais. O capítulo começa com uma breve introdução à fonética e ao processo de produção de fala, seguindo-se a descrição dos órgãos intervenientes no processo de produção dos sons nasais (secção 2.1.2). A secção 2.2 contém informação sobre o sistema vocálico, sendo as características acústicas e de percepção das vogais nasais apresentadas na secção 2.3. Por fim, são referidos alguns estudos sobre estes sons nasais, referindo-se ainda algumas análises recentes efectuadas para as vogais nasais do PE na secção 2.4.2.

2.1 A Nasalidade

2.1.1 Processo de Produção de Fala

O mecanismo de comunicação verbal é um conjunto de procedimentos complexos que envolve diferentes níveis de codificação/descodificação e cujo controlo é, fundamentalmente, assegurado pelo sistema nervoso central.

A ciência que estuda os sons da fala e o modo como são produzidos e percebidos pelos locutores/auditores é a Fonética, que costuma subdividir-se em três áreas de estudo, sendo

¹Utilizamos o alfabeto *Speech Assessment Methods Phonetic Alphabet* (SAMPA) para a transcrição fonética (<http://www.phon.ucl.ac.uk/home/sampa/home.htm>).

²Este capítulo é uma evolução de parte do capítulo 2, entitulado “Preliminares”, da tese de Doutoramento de Teixeira (2000).

estas:

- Fonética articulatória - estudo da produção dos sons pelos órgãos de fala (aparelho articulatório ou fonador);
- Fonética acústica - estudo da estrutura física dos sons, da onda sonora;
- Fonética perceptiva - estudo da recepção e integração dos sons pelos órgãos da audição e diferentes zonas cerebrais.

Na realidade, não existem órgãos especificamente atribuídos à produção de fala, sendo o aparelho fonador uma adaptação secundária dos aparelhos respiratório e digestivo (Figura 2.1).

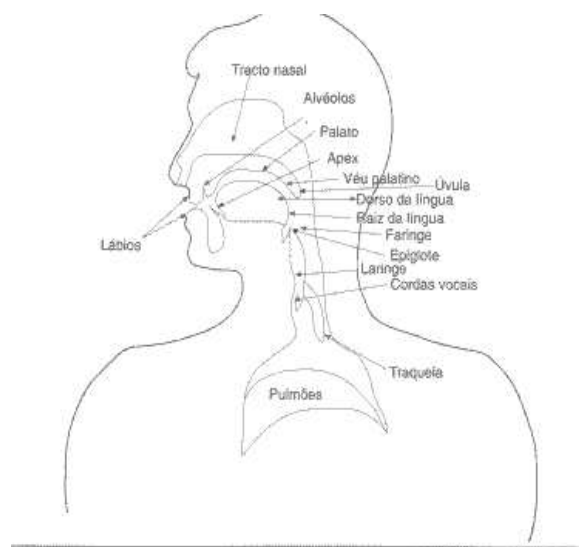


Figura 2.1: Aparelho fonador.

Deste modo, os órgãos associados à respiração e à fala podem ser divididos em três grupos (Moutinho, 2000):

- sistema sub-laríngeo, que inclui os pulmões, os brônquios e a traqueia, que fornece a corrente de ar;
- laringe, onde se encontram as cordas vocais e onde é produzida a energia sonora que se utiliza durante a fala;
- cavidades supra-laríngeas, onde se incluem o tracto vocal e o tracto nasal, que têm como função servir de caixa de ressonância.

O processo de produção de voz começa quando o ar que é expelido pelos pulmões atravessa a traqueia e chega à laringe, onde se encontram as cordas vocais. As cordas vocais são, no

fundo, duas membranas compostas por uma série de ligamentos e músculos, que em determinadas condições podem vibrar, afastando-se e aproximando-se uma da outra. Quando a pressão diminui, as cordas tendem a juntar-se, dando origem a um novo ciclo.

A laringe revela-se, por isso, um órgão fundamental à produção de voz e é o seu funcionamento que fornece o critério de sonoridade. Assim, se as cordas vocais vibram, produzem-se sons vozeados ou sonoros (p.ex. [b,d,g]). Quando estas se encontram afastadas, o ar passa pela laringe sem sofrer vibrações, dando origem a sons não vozeados ou surdos, como [p,t,k].

Nas cavidades supra-laríngeas estão incluídos o tracto vocal e o tracto nasal. No primeiro situam-se as estruturas anatómicas utilizadas na produção de som, os articuladores. Os traços articulatorios com que distinguimos os sons da fala têm origem nesses pontos de articulação e são os seguintes:

- lábio superior
- dentes
- alvéolos
- pré-palato
- palato (palato duro)
- véu palatino (palato mole ou “velum”)
- úvula

Uma corrente de ar pode ser expelida de uma de três formas. A mais comum, nas línguas românicas e noutras, acontece quando a corrente de ar escapa apenas pela cavidade oral - faríngea, produzindo desta forma sons orais.

No entanto, é possível que as cavidades nasais estejam envolvidas neste processo, resultando assim que a corrente de ar seja expelida, total, ou parcialmente através das últimas. Tal acontece quando existe uma passagem através das cavidades orais, mas também através das câmaras nasais, de modo, que o ar passa simultaneamente através de ambos os canais. Assim, os sons objecto deste estudo, isto é os sons nasais, são produzidos quando o véu palatino se encontra abaixado, a úvula se afasta da parede faríngea e o ar passa, simultaneamente pela cavidade bucal e pela cavidade nasal.

Durante esta apresentação não vai ser criada uma oposição formal entre sons onde o ar é expelido parcialmente pelos canais nasais daqueles em que existe uma total utilização destas cavidades. Apenas serão referidos como “nasais” quaisquer tipos de sons em cuja produção participam as cavidades nasais.

Nas secções seguintes é analisada a cavidade nasal, bem como todo o processo de produção de vogais nasais.

2.1.2 Anatomia e Fisiologia das Cavidades Nasais

Os foneticistas reconhecem três câmaras de ressonância no tracto vocal supra-glotal. Duas estão fortemente relacionadas e têm efectivamente um formato de trompa em perfeita continuidade. Estas são as cavidades orais e faríngeas. Acima deste par de cavidades encontram-se as cavidades nasais, mais precisamente, entre o véu palatino e as narinas, por um lado, e entre o palato e o cérebro, por outro. Ambas as cavidades estão ligadas por uma estreita abertura, a abertura velo-faríngea, cujo tamanho é regulado pela movimentação do véu palatino (palato mole).

Na zona do véu palatino, o tracto nasal é designado por nasofaringe, sendo a continuação da faringe. A nasofaringe termina na coana, onde ocorre uma bifurcação iniciando-se duas passagens, laterais, separadas pelo septo nasal.

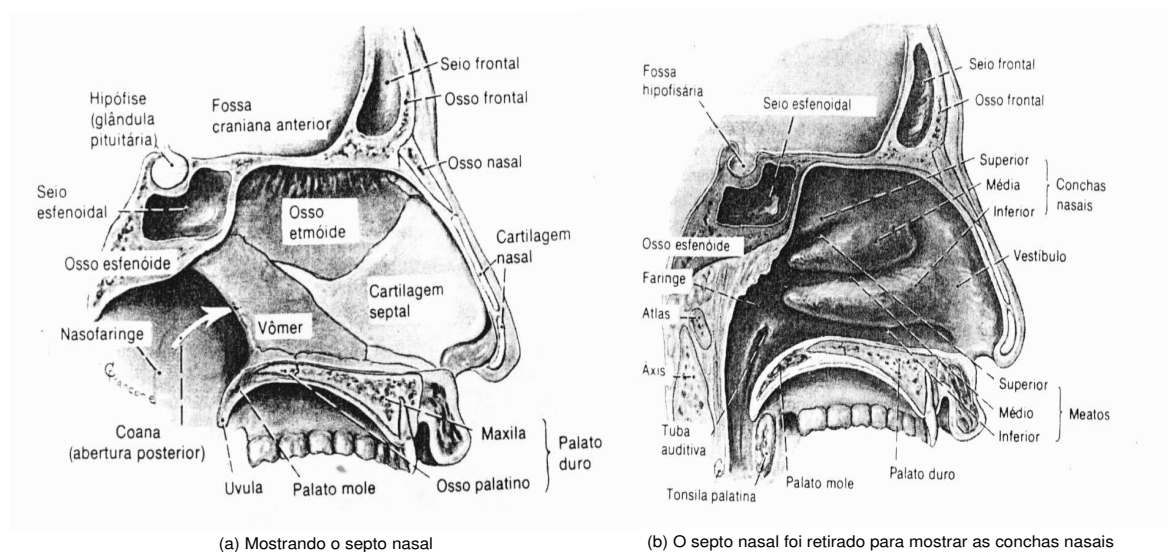


Figura 2.2: Corte sagital da cavidade nasal.

Não existe nas cavidades nasais qualquer órgão móvel, logo não é possível aos oradores modificar o volume e a forma destas câmaras. Assim, as cavidades nasais apenas acrescentam o timbre próprio da sua ressonância aos sons, mas não permitem a produção de diferentes sons nasais. A variedade dos sons nasais é obtida por variações na cavidade bucal.

As dimensões das cavidades nasais variam, no entanto, entre indivíduos, dependendo da sua idade, género e da sua caracterização física.

Em comunicação com a cavidade nasal, através de pequenas aberturas, encontram-se os seios paranasais (Figuras 2.2 e 2.3). Os seios paranasais estão divididos em quatro grupos (maxilares, frontais, esfenoidais e células etmoidais), que ocorrem geralmente aos pares, não sendo, no entanto, simétricos. A principal função dos seios paranasais é a de tornar os ossos do crânio mais leves. Têm também funções secundárias, como a de fornecimento de muco para a cavidade nasal e a de actuarem como caixa de ressonância na produção do som. O seu

papel na produção de sons nasais é ainda alvo de controvérsia, sendo para alguns importante no aparecimento das ressonâncias de frequência mais baixa em sons nasais.

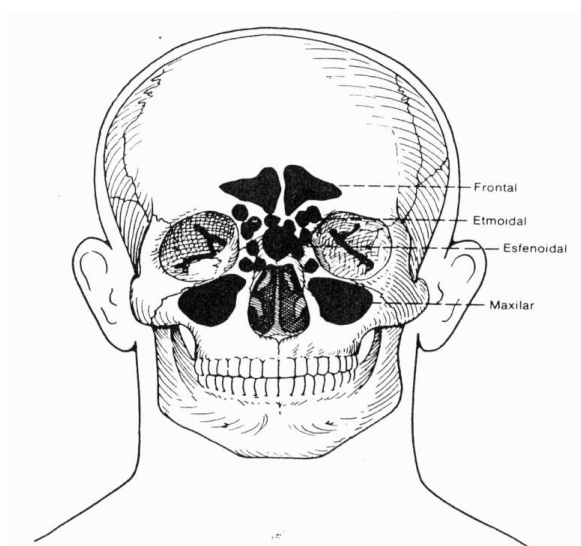


Figura 2.3: Seios Paranasais.

Véu Palatino

O acesso às cavidades nasais da corrente de ar necessária para produzir os sons nasais, é controlado pela posição do velo. O véu palatino, ou velo, pode ser considerado como a continuação flexível do palato duro e é constituído por uma fina camada de fibras musculares, tecido, vasos sanguíneos, nervos e glândulas, o que o torna flexível permitindo que este possa ser subido ou descido segundo o controlo do orador, possibilitando, assim, a passagem do ar pelas cavidades nasais.

Os movimentos do véu palatino são controlados por diversos músculos, que trabalham em cooperação. A subida do velo é conseguida principalmente pela acção de dois músculos, o *levator veli palatini* e o *superior pharyngeal constrictor*. O primeiro liga a superfície frontal do velo à base do crânio e sobe e retrai o velo cerca de 2 cm, em média para um adulto. O segundo, que tem por função principal empurrar a comida para baixo em direcção ao esófago, pode também contrair para subir o velo. Outro músculo, o *tensor palatini*, também actua para alongar e tornar tenso o velo ao subir. A descida é afectada por dois músculos, assistidos em parte pelo relaxamento dos músculos intervenientes na subida e pela força da gravidade. Por um lado, existe o *palatoglossus* que se estende para baixo, desde a parte inferior da superfície do velo, dividindo-se em dois e depois ligando-se à língua. Este tanto pode baixar o velo, como subir a parte posterior da língua. Por outro lado, existe o *palatopharyngeous*, um músculo longo e fino, que liga o velo e a parte posterior da cartilagem tiróide e da parede lateral da faringe. Apesar de todos os indivíduos possuírem o mesmo conjunto de músculos, não os operam necessariamente de forma igual para atingir um mesmo movimento durante a produção de voz. Por exemplo, o *palatoglossus* pode não ser activado por certos indivíduos

para causar o abaixamento do velo.

2.2 O Sistema Vocálico Nasal do Português

Em PE é geralmente referida a existência de cinco vogais nasais (Sampson, 1999) , sendo estas: o [6̃] (**canto**), [ẽ] (**tenta**), [ĩ] (**pinta**), [õ] (**ponte**), e [ũ] (**fundo**). É reportada, ainda, a existência de outras três vogais, o [Ẽ], o [Õ] e o [ã]. O [ã], resultante da crase de [6] com [6̃], como em *a andar* e os [Ẽ] e [Õ], como, por exemplo, nas palavras *vende* e *ontem*, têm se vindo a generalizar cada vez mais, pelo menos nestes termos, no Português dito padrão.

Existem também ditongos nasais, como por exemplo, [w6̃] (**quando**), [wẽ] (**aguentar**), [j6̃] (**fiando**), [wĩ] (**ruim**), [6̃j] (**quem**), [õj] (**põe**), [ũj] (**muito**), [6̃w̃] (**pão**). Como exemplo de um tritongo nasal temos o [w6̃w̃] (**enxaguam**).

Segundo diversos estudos de fonética (Stevens *et al.*, 1987), as vogais nasais do Português Europeu diferem das vogais nasais de outras línguas, como o Francês, devido a serem fortemente nasalizadas apenas junto ao fim .

2.3 Propriedades Acústicas das Vogais Nasais

Acusticamente, as vogais surgem da ressonância nas cavidades oral-faríngea. A fonte de energia para a ressonância é a excitação glotal, causada pela vibração periódica das cordas vocais. Associado a cada vogal existe um conjunto de formantes, ou ressonâncias, localizadas em frequências específicas. As diferenças para os valores das formantes são determinados pelo tamanho e forma da câmara de ar na cavidade oral-faríngea. Esta configuração pode ser prontamente modificada, principalmente como resultado do movimento da língua (levantamento/abaixamento; avanço/retracção) mas também pelo movimento dos lábios e maxilas. Por exemplo, o aumento da altura da língua está fortemente correlacionado com o decréscimo da frequência de F1 e vice-versa. Também a retracção do corpo da língua está correlacionada com a diminuição da frequência de F2 (Sampson, 1999).

A adição da cavidade nasal complica a situação, não só ao adicionar novas ressonâncias, mas também ao originar anti-ressonâncias, estas normalmente designadas por pólos e zeros, respectivamente. Os zeros ou anti-ressonâncias servem para amortecer ou mesmo cancelar qualquer ressonância que possa estar situada na vizinhança da sua frequência. Quando o canal nasal está fechado, algum som pode propagar para as cavidades nasais, mas como as cavidades estão bloqueadas, os zeros e pólos associados coincidem na frequência e cancelam-se mutuamente. No entanto, quando este canal está aberto existe um acoplamento nasal, e os zeros e pólos criados nas cavidades nasais tornam-se diferenciados na frequência. Quanto maior for o nível de acoplamento nasal, ou seja, quanto maior for a abertura do canal nasal, mais estão os pares de pólos e zeros separados na frequência.

Quais são então as propriedades acústicas das vogais nasais? Foi feita uma investigação exhaustiva das características espectrais das vogais nasais ao longo das últimas quatro décadas que originou resultados diversificados e por vezes mesmo contraditórios entre si. Existem, no entanto, alguns pontos de acordo. As marcas principais de nasalidade de uma vogal parecem ser: a modificação do espectro nas baixas frequências, particularmente na vizinhança da primeira formante oral; a existência de uma formante nasal cerca dos 250 Hz (característica também das consoantes nasais); a existência de um zero que interage com a primeira formante oral, reduzindo a sua amplitude e aumentando a sua largura de banda e a modificação do espectro nas frequências mais elevadas, resultando numa distribuição mais difusa da energia (Sampson, 1999, p. 7).

No entanto, é necessário considerar que a nasalidade em vogais pode variar significativamente na forma como é realizada fisicamente, de idioma para idioma, tornando qualquer conclusão universal sobre as suas propriedades acústicas bastante problemática.

2.3.1 Percepção de Vogais Nasais

Como foi visto, ocorrem complexas modificações acústicas quando a nasalidade é adicionada à articulação das vogais. Assim, e uma vez que é com estes sinais complexos que o orador é confrontado e tem de interpretar, surge um conjunto de questões que interessa analisar (Sampson, 1999, p. 9):

1. Qual a propriedade - ou conjunto de propriedades - que permite a um ouvinte perceber uma vogal como sendo nasal?
2. A percepção da nasalidade é independente da vogal e da língua?
3. Que outros aspectos da qualidade da vogal são influenciados pela nasalização ?

Propriedades influenciadoras da percepção de nasalidade

Tal como referido anteriormente, uma vogal é rapidamente percebida como nasal se a proeminência na área da primeira formante é reduzida e a largura de banda aumenta. De facto, já na década de 1950 foi mostrado por Delattre (1954), com a ajuda de síntese, que a nasalidade pode ser detectada apenas pela redução da intensidade da primeira formante de uma vogal oral. O estudo realizado por Hawkins e Stevens (1985), usando voz sintética apresentada a falantes de Gujarati, Hindi, Bengali, e Inglês, aponta para que a redução da proeminência e aumento da largura de banda, na zona da primeira formante, sejam independentes da língua considerada.

Outras propriedades espectrais, temporais e contextuais, parecem também influenciar a percepção de nasalidade.

Uma é o aumento progressivo do nível de nasalidade durante a articulação de uma vogal nasal. Reenen, num trabalho de 1982 (citado em Sampson, 1999, p. 10) atribui muita

importância a esta característica dinâmica, vendo-a como crucial para a percepção da nasalidade numa vogal, independentemente da língua. Uma experiência realizada por Linthorst, em 1983 (referido em Sampson, 1999, p. 10), fornece também suporte para a importância desta propriedade na percepção. Foram gravadas três palavras francesas *même*, *dais*, *baie*, extraídas as vogais e, depois, adicionadas, em várias combinações, às consoantes iniciais das palavras usadas. Representando a vogal parcialmente nasal da primeira palavra por [E] e a vogal oral das outras duas por [ε], os nove ouvintes franceses perceberam quase sempre a sequência [εE] como vogal nasal, e as outras combinações, incluindo o [E] e [EE], como vogal oral. Não foram realizados estudos similares para outras línguas.

Evidências fonológicas da relação entre a duração das vogais e nasalização motivaram experiências em que se manipulou a duração para determinar os efeitos na nasalização das vogais. Delattre e Monnot em 1968 (citados em Teixeira, 2000), usando o *Pattern Playback*, sintetizaram nove versões de uma sequência CVC, diferindo apenas na duração da vogal. Ouvintes franceses e americanos identificaram as vogais mais breves como sendo orais e as vogais mais longas como nasais. Neste estudo, o grau de nasalização foi mantido constante e igual a um valor intermédio entre o de uma vogal oral e a vogal nasal.

Dois tipos de resultados experimentais sugerem uma relação entre a altura da vogal e a percepção de nasalidade. Primeiro, estudos usando voz natural descobriram que vogais baixas têm uma maior tendência a serem percebidas como nasais. Segundo, estudos com voz sintetizada mostram que vogais baixas requerem um maior acoplamento nasal para serem percebidas como nasais. Esta aparente contradição deve ter origem na utilização de diferentes aberturas do velo. Em voz natural, as vogais baixas tendem a ser produzidas com uma posição do velo mais baixa mesmo em contextos orais.

A percepção da nasalidade de uma vogal é também influenciada pelo contexto fonético em que a vogal ocorre. Por exemplo, ouvintes americanos julgam vogais nasais como mais nasais quando em contextos não nasais (Beddor, 1983, p. 179). Uma possível explicação para este resultado é a de que o conhecimento pelos ouvintes dos efeitos de coarticulação leva-os a atribuir a nasalidade às consoantes nasais adjacentes, ouvindo vogais nasais, num contexto nasal, como não nasais.

2.4 Estudos de Vogais Nasais

A nasalidade em vogais foi estudada por muitos investigadores usando as mais variadas técnicas. Referem-se nesta secção alguns estudos representativos das técnicas utilizadas, sem se pretender, no entanto, efectuar um levantamento exaustivo. Não se inclui, em geral, informação acerca dos resultados dos trabalhos citados, pois já foram referidos nas secções anteriores. Os estudos das vogais nasais do Português são tratados em separado.

2.4.1 Técnicas Utilizadas

Os primeiros estudos das vogais nasais foram realizados utilizando técnicas de análise do sinal de voz.

O estudo do espectro de vogais nasais, nomeadamente o comportamento dos pólos e zeros da função de transferência para vogais nasais, foi efectuado por Fujimura (1960), utilizando um modelo simplificado das cavidades oral e nasal. Este utilizou um método gráfico para obtenção das ressonâncias e anti-ressonâncias e, de modo a tornar o problema tratável, não considerou as perdas no seu modelo.

A medição directa da função de transferência para vogais nasais, em seres humanos, foi efectuada por Båvegård *et al.* (1993) usando uma técnica designada de *sweeptone*. Colocaram uma fonte de vibração no pescoço, muito próximo da posição da laringe, e mediram a radiação dos lábios e narinas em conjunto. Variando a frequência de oscilação da fonte de vibração, mediram a função de transferência entre os 100 e os 5000 Hz.

Diversos estudos utilizaram sinal natural editado. Num exemplo deste tipo de estudos, que pretende testar a hipótese de que a nasalização numa vogal nasal deve ser mais evidente perceptualmente, ao atenuarem-se as consoantes nasais adjacentes, Kawasaki (1986) gravou três sílabas constituídas pela nasal [m] e uma vogal, atenuando de seguida as consoantes nasais. Os estímulos assim obtidos foram utilizados num teste perceptual, tendo-se constatado resultados que suportam aquela hipótese.

O movimento do velo durante a produção de vogais nasais foi investigado através de técnicas como: o *velotrace*, endoscopia, EMMA. Uma apresentação das várias técnicas instrumentais utilizadas em estudos de sons nasais pode ser encontrada em (Huffman e Krakow, 1993).

Uma parte substancial dos estudos acerca de vogais nasais consistiu na realização de testes perceptuais utilizando estímulos obtidos usando técnicas de síntese. Podem dividir-se estes estudos em dois tipos: um, manipulando directamente as características espectrais do som, o outro, utilizando síntese articulatória.

Os primeiros estudos baseados na manipulação directa do espectro foram realizados por Delattre (1954), usando o *Pattern Playback* desenvolvido nos Laboratórios Haskins. Este conseguiu obter vogais nasais, adicionando energia na vizinhança da frequência fundamental. Takeuchi *et al.* (1975) adicionaram pares de pólos e zeros ao espectro de voz natural a várias frequências, descobrindo que os sons eram considerados mais vezes como nasais quando a adição era feita na vizinhança da primeira formante. Vários estudos utilizaram o sintetizador de formantes para o estudo das vogais nasais.

Diversos estudos utilizaram síntese articulatória para, através de simulação, obter a função de transferência, respectivas formantes e anti-formantes, e/ou para obter os estímulos necessários para a realização de testes perceptuais. O primeiro foi o realizado por House e Stevens (1956), usando o sintetizador articulatório desenvolvido por Stevens *et al.* (1953). Estes estudaram o efeito nas primeiras formantes de diferentes aberturas do velo e realizaram

testes perceptuais usando estímulos produzidos pelo sintetizador. Fant (1960) estudou para várias vogais nasais o efeito no espectro de diversos factores, como por exemplo a área de acoplamento, utilizando para tal um sintetizador analógico. Um estudo relevante foi realizado por Krakow *et al.* (1988). Estudaram a influência do contexto na percepção de vogais nasais, comparando a sua percepção na presença e ausência de uma consoante nasal adjacente. Mais estudos foram, entretanto realizados.

Desde 1995 que se tem vindo a trabalhar, no Departamento de Electrónica e Telecomunicações da Universidade de Aveiro, na síntese dos sons nasais do Português (Teixeira, 2000). Foram já efectuados alguns testes perceptuais, para as vogais nasais do PE, utilizando estímulos produzidos pelo sintetizador em desenvolvimento nesta instituição. Os resultados indicam que vogais sintetizadas com a mesma configuração para a língua são percebidas de forma diferente por falantes nativos do PE quando é acrescentada nasalidade (Teixeira *et al.*, 2003b).

2.4.2 Estudos das Vogais Nasais Portuguesas

Diversos estudos, usando metodologias diversas e com objectivos também bastante diferentes, foram já realizados acerca das vogais nasais do Português. Resumem-se, de seguida, algumas contribuições importantes para o conhecimento dos sons nasais vocálicos do Português, particularmente na sua variante europeia.

Os primeiros estudos foram realizados usando, apenas, a capacidade auditiva dos investigadores. Viana (1883) refere a diferença das vogais nasais do Português em relação às do Francês. Segundo ele, a nasalização em Português não é acompanhada de guturalização; o timbre das vogais nasais é o mesmo das vogais orais e não há em Português nenhuma vogal nasal equivalente, em timbre, a qualquer vogal nasal francesa. Considera que a nasalidade portuguesa é mais fraca do que a francesa. O mesmo autor refere, ainda, que apenas existem vogais e ditongos nasais puros antes de repouso, de uma vogal ou de uma consoante contínua. Quando a uma vogal se segue uma explosiva, além dessa vogal nasal, ouve-se, atenuada, uma consoante nasal homorgânica com essa explosiva.

Os estudos de Louro (1954-1955) (citado em Lacerda e Head, 1966), usando valores de formantes e radiografias, levam-no a considerar as vogais nasais como sons oro-nasais, visto que a ressonância nasal se junta à vogal emitida normalmente pela boca. Para ele apenas o [a~] pode ser oro-nasal ou exclusivamente nasal, neste caso apenas emitido pelas fossas nasais. As outras vogais, quando mediais ou no interior das frases, são geralmente ligadas (ou mesmo substituídas na sua parte final) por um [a~] (formando com elas uma espécie de ditongo decrescente). Para Louro são as vibrações deste [a~] que fazem pensar na existência de verdadeiras consoantes nasais, em fim de sílaba interna, em Português. Relativamente a essas consoantes nasais finais, antes de oclusiva, apesar de serem representadas na actual ortografia, não existe a consciência de se pronunciarem, não lhes correspondendo quaisquer movimentos articulatorios activos próprios. Especialmente no caso da nasal ser seguida de consoante oclusiva labial, pode surgir o equivalente fonético da consoante nasal, do *m*. Mas

este *m* é também inconsciente e passivo, resultando apenas do mesmo [@ ~] da oclusão da boca para a pronúncia da labial seguinte.

Lacerda e Strevens, (citados em Lacerda e Head, 1966, p. 9), em 1956, utilizaram um extensor sonoro (*speech-stretcher*) e verificaram que as vogais nasais em Português têm um segmento inicial cujo grau de nasalidade varia grandemente. Em alguns casos a nasalidade pode ser tão diminuta que na prática se pode considerar inexistente, considerando-se o início como oral. A natureza deste segmento inicial é dependente do contexto, havendo contextos que favorecem mais a existência de um início oral do que outros.

Para Barbosa (1961), antes de uma pausa, a vogal nasal é em princípio pura, podendo, no entanto, existir na fase final uma oclusão da passagem oral na região velar. Para este autor, essa consoante nasal passa despercebida a ouvidos não treinados em análise fonética, apesar de ser bem notória a sua presença em espectrogramas. Antes de oclusiva, considera a existência de uma consoante nasal nítida e audível, com ponto de articulação dependente da consoante seguinte. Barbosa refere nunca ter encontrado exemplos de vogal nasal seguida de oclusiva oral sem a existência de consoante nasal intermédia.

Em 1964, Head (citado em Almeida, 1976, p. 357) estudou as características acústicas das vogais nasais portuguesas usando espectrogramas. Refere uma maior densidade de formantes nas vogais nasais do que nas correspondentes vogais orais. Para a maioria das vogais, /*ĩ*, *ẽ*, *õ*, *ũ*/, existe, segundo Head, uma formante adicional entre as duas primeiras; para o /*6*~/ a formante adicional aparece abaixo da posição da formante mais baixa do /*a*/.

Um estudo instrumental, usando o denominado pneumocromográfico, foi efectuado por Lacerda e Head (1966), o que lhes permitiu ter informação simultânea da radiação nasal e da radiação oral em traçados separados. Foi efectuado o estudo da nasalidade de vogais nasais em posição final, antes de oclusiva e antes de restritiva. Como resultado das suas análises obtiveram que as vogais nasais são sempre parcialmente orais, com nasalidade médio-final ou final, excepto no caso de ocorrer uma consoante nasal anterior. Neste caso, são inteiramente nasais, isto é, com radiação nasal desde o início até ao final.

A análise acústica efectuada por Martins em 1973 (citada em Mateus, 1975, p. 94) distinguiu nas vogais de palavras como *canto*, *campo*, uma primeira área sem ressonância nasal, seguida de uma área em que a ressonância nasal aumenta.

Stevens *et al.* (1987) efectuaram testes perceptuais usando um conjunto de estímulos gerados por um sintetizador de formantes. Os estímulos foram obtidos variando a duração da vogal, a nasalização e a duração do murmúrio nasal que segue a vogal. Foi pedido a falantes de Português, Inglês e Francês, para:

1. assinalarem a presença ou ausência de nasalização;
2. classificarem o estímulo em relação à sua naturalidade.

Foi estudado, apenas, o caso de vogal nasal entre duas oclusivas, o [t].

Os falantes de Português preferiram estímulos em que existia murmúrio nasal e mais

nasalização. Apesar de preferirem mais nasalização, não era necessário que a nasalização ocorresse durante toda a realização da vogal.

No IV encontro nacional da Associação Portuguesa de Linguística, Drenska (1988) faz uma análise acústica comparativa entre as vogais nasais do Português e do Búlgaro. Esta conclui pela manifestação da nasalidade nas vogais num plano temporal, indicando que a parte inicial do segmento pode ser considerada como oral e que, apenas num pequeno número de casos, a fase final das vogais nasais portuguesas é articulada ao segmento consonântico homorgânico da consoante seguinte, sendo este segmento de pequena duração.

Silva (1995) efectuou a análise de vogais nasais em vários contextos, utilizando o método *Recursive Least Squares* (RLS) para obtenção dos parâmetros de um modelo auto-regressivo. Obteve dados importantes acerca da duração e da variação ao longo da realização das vogais nasais das primeiras três formantes, respectivas amplitudes e larguras de banda. As suas análises levaram-no a concluir pela existência de dois estados estáveis, além de uma zona de transição. Silva considera que a pronúncia das vogais nasais, em Português, é efectuada da seguinte forma:

- Posicionamento dos articuladores para a pronúncia da vogal oral correspondente. Esta pronúncia terá entre 60 e 100 ms de duração;
- Abertura da passagem velo-faríngea, dando origem a uma fase de transição, com a duração de 30 a 50 ms;
- Aparecimento de uma segunda fase estável, correspondente à situação de acoplamento do tracto nasal. Esta fase possui características muito semelhantes para as várias vogais nasais.

Galvão (1998) efectuou medidas de *nasalance*³ de sete falantes do Português, obtendo confirmação para a existência de um murmúrio nasal ou traço de uma consoante nasal, na fase final de vogais nasais, quando situadas antes de consoantes oclusivas. Confirmou, também, que vogais elevadas possuem uma maior percentagem de *nasalance* do que vogais médias.

Para além dos estudos da Fonética, outros estudos poderão ser úteis para a compreensão das vogais nasais. Um tipo de estudo com interesse é o da forma como se deu o aparecimento e a evolução deste tipo de sons em Português (Lacerda e Head (1966); Sampson (1999)) e nas línguas românicas em geral (Sampson (1999)).

Síntese articulatória dos sons nasais do Português

Desde 1995 que se tem vindo a trabalhar, no Departamento de Electrónica e Telecomunicações da Universidade de Aveiro, na síntese dos sons nasais do Português (Teixeira, 2000). A principal motivação prende-se com a obtenção de síntese com qualidade próxima da natural.

³*Nasalance* é o quociente da energia nasal pela energia total, $N/(N + O)$

Os resultados das experiências efectuadas mostram que a variação do velo, e mesmo de outros articuladores, influencia a produção e percepção da nasalidade (Teixeira *et al.*, 2003b).

A síntese articulatória permite realizar simulações e, funcionando como um falante versátil, obter estímulos para a realização de testes perceptuais, sendo útil em estudos indispensáveis na obtenção de novos conhecimentos sobre os mecanismos de produção e percepção.

Encontra-se em desenvolvimento um sintetizador articulatório especialmente vocacionado para a síntese do Português. Tendo por objectivo melhorar o conhecimento acerca dos sons nasais, pesquisa-se a interacção entre a fonte de excitação glotal e as cavidades acima da laringe, e o efeito da dinâmica dos articuladores.

O Projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português”

O projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português” foi motivado pela existência de uma investigação multidisciplinar entre linguistas e engenheiros (Departamento de Línguas e Culturas - CLC e Departamento de Electrónica e Telecomunicações - IEETA da Universidade de Aveiro), com o objectivo de estudar e aplicar as técnicas da síntese articulatória ao estudo e síntese do PE.

Foi referido anteriormente, que desde 1995 que se tem vindo a desenvolver, na Universidade de Aveiro, pesquisa no domínio da síntese dos sons nasais do PE. No entanto, não existia um estudo acústico detalhado, considerando a variação das características ao longo da produção das vogais nasais. Por exemplo, informação sobre o comportamento da fonte glotal durante a produção dessas vogais, bem como informação acerca da sua variação regional, não estava disponível, tornando imprescindível a obtenção de dados sobre a forma como estes sons são naturalmente produzidos.

Assim, neste projecto utilizou-se o EGG para obter informações sobre a excitação glotal ocorrida no momento de produção das nasais do Português. O EGG, de simples utilização e não interferindo com o sinal de fala, revela-se um instrumento de grande utilidade numa pesquisa deste tipo e indispensável para a análise da variação sincrónica do pitch.

Foram já publicados alguns estudos utilizando este *corpus* (Teixeira *et al.* (2003b) e A.Teixeira *et al.* (2004a)) que contemplam a diversidade contextual e regional que este possui.

Capítulo 3

Corpora

O estudo detalhado das vogais nasais do PE exige a utilização de grandes quantidades de dados, que abranjam áreas relevantes, como, por exemplo, a variação regional e contextual e o comportamento da fonte glotal durante a produção dos sons nasais.

Deste modo, a existência de um conjunto suficientemente extenso de dados linguísticos é imprescindível para a obtenção de resultados significativos. No entanto, tarefas como a escolha do conteúdo de um *corpus*, a escolha da fonte, a recolha, e até a anotação dessas grandes quantidades de material, não são simples.

A apresentação do material a gravar ao locutor é outro dos parâmetros a ter em conta. Caso exista necessidade de controlar o material gravado, de forma a que seja pronunciado o mais aproximadamente possível do que é pretendido, é indicada a utilização de fala lida. Caso se pretenda estudar fala natural, e a presença de pausas, hesitações, ruídos ou sobreposições (que implicam, em geral, uma segmentação manual) não influenciem a análise pretendida, é aconselhada a gravação de fala espontânea.

A escolha dos locutores, bem como o que diz respeito às condições de gravação, em particular, às condições acústicas, são também princípios a considerar. *Corpora* recolhidos com base em inquéritos no terreno, em oposição à utilização de cabinas insonorizadas, são escolhas cruciais para a obtenção dos dados pretendidos.

O material utilizado neste trabalho provém de dois *corpora*, recolhidos no âmbito dos projectos de pesquisa realizados na Universidade de Aveiro, referidos na secção 2.4.2. A descrição detalhada destes *corpora* será efectuada nas secções seguintes.

3.1 *Corpus* EMMA

O sintetizador articulatório, em desenvolvimento na Universidade de Aveiro (secção 2.4.2), necessita de informação sobre os movimentos dos articuladores durante a produção de sons nasais, em particular, de informação sobre a posição da língua, lábios e véu palatino. Com este objectivo, foi recolhido, através de um sistema EMMA, um *corpus* de aproximadamente

uma hora, que é apresentado nas secções seguintes.

O acrónimo EMMA provém da designação inglesa “*ElectroMagnetic Midsagittal Articulography*”, nome proposto por J. Perkell, em 1992, para o sistema desenvolvido no *Massachusetts Institute of Technology* (MIT). Esta técnica pertence à categoria dos dispositivos que permitem obter dados sobre as trajectórias de pontos dos articuladores num espaço bidimensional. A tecnologia EMMA é muito adequada a estudos de coarticulação pois permite recolher uma grande variedade de frases e/ou palavras numa única sessão.

A utilização desta tecnologia permite monitorizar os movimentos no plano médio-sagital da maioria dos articuladores, como o lábio inferior, maxilar, língua e velo. É, no entanto, de maior interesse para a língua, visto que o lábio inferior e o maxilar são de fácil acesso a outros sistemas. Já o velo coloca dificuldades acrescidas de fixação dos sensores, havendo, por isso, poucos dados acerca do seu comportamento, obtidos através desta técnica.

3.1.1 Constituição do *Corpus*

O *corpus* foi elaborado com base nas questões apresentadas de seguida, que foram organizadas em 3 grandes grupos (Teixeira e Vaz, 2001). No primeiro incluem-se as perguntas mais relevantes para o estudo. O segundo grupo tem relevância intermédia e o terceiro é de importância relativa na pesquisa.

- Grupo 1
 - Como varia a posição do velo em vogais nasais pronunciadas entre duas oclusivas? Qual a duração da parte oral inicial da vogal nasal? Qual a velocidade de abertura-fecho do velo? Qual a duração da consoante nasal criada por coarticulação no final da vogal nasal?
 - Como se movimenta o velo em sequências consoante nasal seguida de vogal nasal?
 - Qual a posição da língua em vogais nasais e orais correspondentes (ex. [i] versus [ĩ])?
- Grupo 2
 - Em posição final, as vogais nasais [ẽ] e [õ] tomam a forma de ditongos? Em que casos?
 - Como se comporta o velo em sons nasais em final de palavra e/ou frase?
 - Como são as vogais nasais isoladas (em termos de posição do velo e dos outros articuladores)?
 - Qual o comportamento do velo, língua, lábios e maxilar nos ditongos?
- Grupo 3
 - Como é feito o contraste oral/nasal entre consoantes nasais?
 - Que acontece ao velo e articuladores orais em: vogal nasal seguida de vogal oral, vogal nasal seguida de consoante nasal e vogal oral seguida de vogal nasal?
 - Existe nasalidade em palavras como lâ?

3.1.2 Recolha e Informante

O *corpus* foi recolhido em Fevereiro de 2001, no Institut für Phonetik und Sprachliche Kommunikation da Universidade Ludwig-Maximilians, em Munique. Foi utilizado o sistema AG100 comercializado pela empresa alemã Carstens. Na recolha os 10 sensores foram distribuídos tal como ilustrado na Figura 3.1.

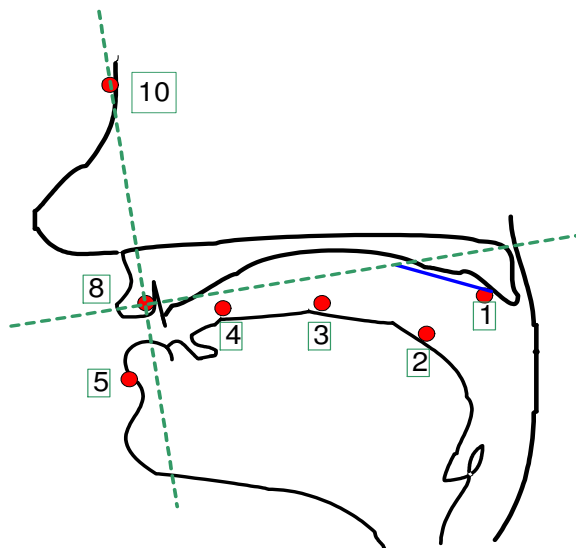


Figura 3.1: Diagrama mostrando a posição dos vários sensores utilizados.

Na figura 3.1 pode verificar-se a posição dos seguintes sensores:

1. vélo, colado numa tira de acetato segura por um palato artificial;
2. parte posterior da língua (*tongue back*). Para não interferir com o sensor do vélo, não foi possível colocar muito atrás;
3. dorso (*tongue dorsum*), colocado aproximadamente no ponto médio entre os sensores 2 e 4;
4. próximo do ápice da língua (*tongue blade*);
5. no lábio inferior;
6. utilizado no final para traçado do palato, não foi fixado durante a recolha do *corpus*;
7. não foi utilizado;
8. para referência, nos incisivos superiores;
9. não foi utilizado;
10. para referência, colocado no nariz.

O informante, um adulto do sexo masculino, recebeu informação através de um ecrã de computador do material a pronunciar, que consistia numa sequência de 3 ou 4 palavras, ou uma frase. Caso fosse detectado algum problema de pronúnciação ou de funcionamento do sistema, efectuava-se uma repetição.

Para além do sinal proveniente dos vários sensores do sistema EMMA, foi gravado o sinal de voz utilizando um sistema *Digital Audio Tape* (DAT). O segundo canal do DAT foi utilizado para gravação de um sinal de sincronismo, utilizado no processo de criação de ficheiros de som sincronizados com os dados EMMA. Toda a sessão foi gravada em vídeo. Os dados foram processados e verificados de forma a serem utilizáveis em MATLAB¹.

Para facilitar a análise, os dados foram anotados usando quatro níveis: palavra, fonemas, eventos relativos ao velo e eventos relativos à passagem oral.

Os resultados das primeiras análises foram apresentados em (Teixeira e Vaz, 2001). Novas análises, efectuadas após uma anotação mais alargada dos ficheiros existentes, foram realizadas durante o trabalho apresentado neste documento (Capítulo 5).

A título de exemplo, apresenta-se, na Figura 3.2, o sinal de voz, o espectograma e as trajetórias do velo e lábio inferior, durante a produção da palavra “amamos”.

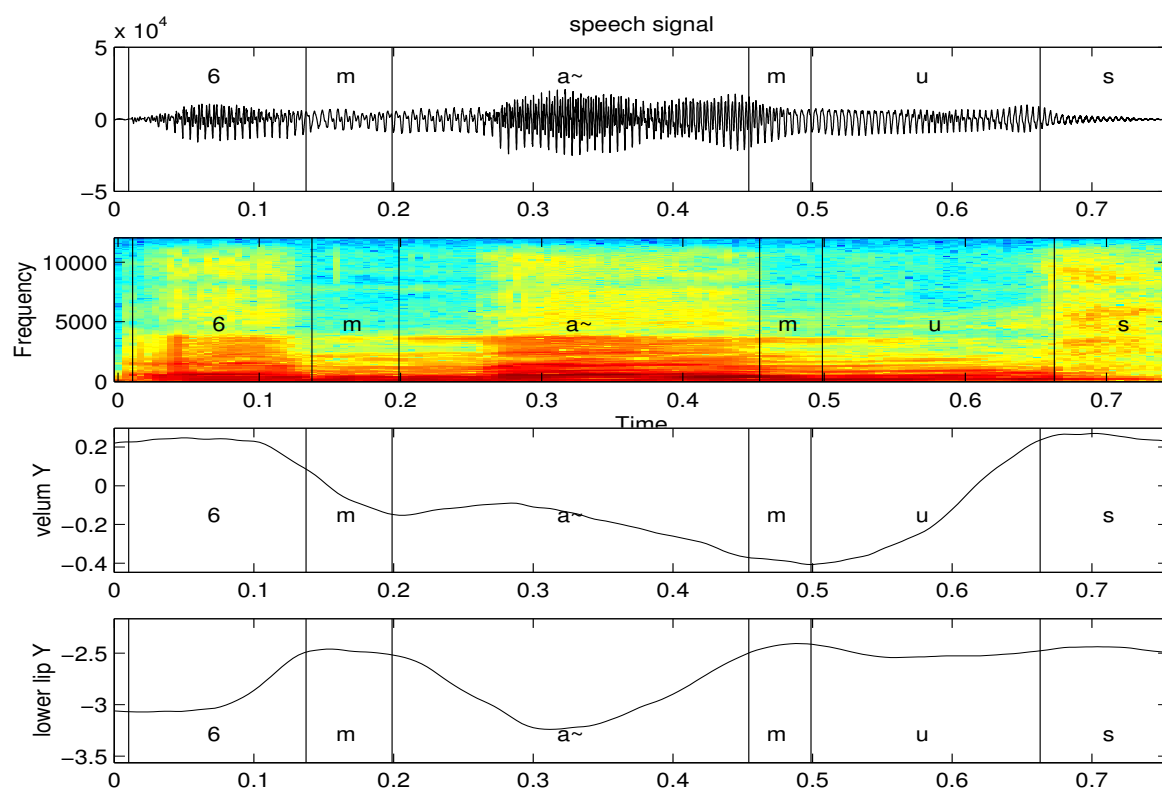


Figura 3.2: Exemplo retirado do *corpus* EMMA.

¹<http://www.mathworks.com>

3.2 *Corpus* Regional

Com o objectivo de explorar áreas pouco desenvolvidas, em particular, as variedades regionais das nasais do PE, foi recolhido, no âmbito do projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português”, um *corpus* que contempla as várias regiões de Portugal Continental (Alentejo, Algarve, Beira Interior, Beira Litoral, Minho e Trás-os-Montes). Foram efectuadas gravações, quer do sinal de voz, quer do sinal proveniente do EGG, que fornece informação sobre o comportamento das cordas vocais. A descrição detalhada deste *corpus* é efectuada nas secções seguintes.

3.2.1 Constituição do *Corpus*

O *corpus* foi organizado de forma a contemplar os vários contextos fonéticos, bem como as diferentes variedades regionais. Existiu, também, a preocupação de incluir todas as vogais nasais do PE, em todos os contextos em que estas podem ocorrer, tal como, palavras contendo consoantes nasais e vogais orais, permitindo deste modo a comparação. A facilidade de representação gráfica, bem como a natureza comum das palavras, foram dois outros requisitos impostos.

De modo a facilitar a organização e gravação, o *corpus* foi dividido em várias tabelas, as quais são sumariadas na Tabela 3.1.

Tabela	Contexto	Palavra	Transcrição	Total
1	entre oclusivas surdas	<i>tinto</i>	[ti~tu]	17
2	entre oclusivas sonoras	<i>gongo</i>	[go~gu]	18
3	após oclusivas surda e antes de oclusivas sonora	<i>tango</i>	[t6~gu]	15
4	após oclusivas sonora e antes de oclusivas surda	<i>banco</i>	[b6~ku]	9
5	entre fricativas	<i>sonso</i>	[so~su]	15
6	palavras com consoantes nasais	<i>monte</i>	[mo~t@]	15
7	antes ou após uma oclusiva	<i>cinto</i>	[si~tu]	15
8	após uma consoante líquida	<i>píncel</i>	[pi~sEl]	9
		<i>renda</i>	[re~d6]	
9	vogais orais em vários contextos	<i>pato</i>	[patu]	30
10	vogais nasais no início ou fim da palavra	<i>entrada</i>	[e~trad6]	9
		<i>rim</i>	[ri~]	
11	sequências contendo vogais nasais nos limites da palavra	<i>lã azul</i>	[l6~6zul]	11
12	ditongos nasais	<i>também</i>	[t6~b6~j~]	7

Tabela 3.1: Os vários contextos contemplados no *corpus* regional.

Foram contemplados contextos para as vogais orais, com o objectivo de confrontar, com os mesmos falantes e nos mesmos contextos, as características das vogais nasais estudadas, com as orais correspondentes.

O *corpus* regional contém, no total, 59 palavras com vogais nasais entre oclusivas (orais), mais 30 palavras contendo vogais orais.

Algumas das palavras seleccionadas para o contexto $C\tilde{V}C$ são apresentadas, a título exemplificativo, na Tabela 3.2. A totalidade das palavras que constituem o *corpus* pode ser consultada no Anexo A.

	[6̃]	[ẽ]	[ĩ]	[õ]	[ũ]
p-p					
p-t	pantufa [p6̃tuf6]	pente [pẽt@]	pintor [pĩtor]	ponte [põt@]	
p-k	pancada [p6̃kad6]	penca [pẽk@]			
t-p	tampa [tãp6]	tempo [tẽpu]			
t-t	estante [St6̃t@]	oitenta [oitẽt6]	tinto [tĩtu]	tonto [tõtu]	
t-k	tanque [t6̃k@]				
k-p	campo [k6̃pu]			compras [kõpt6S]	cumprimentar [kũprimẽtar]
k-t		quente [kẽt@]	quinta [kĩt6]	conta [kõt6]	
k-k				conquistador [kõkiSt6dor]	

Tabela 3.2: Excerto das palavras seleccionadas para o contexto vogal nasal entre oclusivas não vozeadas.

Na recolha foi possível obter, simultaneamente, o sinal de fala e reunir informação objectiva sobre o comportamento das cordas vocais através da utilização do EGG, uma tecnologia não invasiva. Esta dupla recolha é desejável e necessária para uma obtenção mais fiável da frequência fundamental e identificação dos períodos de fecho das cordas vocais. A Figura 3.3 apresenta o EGG e o sinal proveniente deste (traçado inferior).

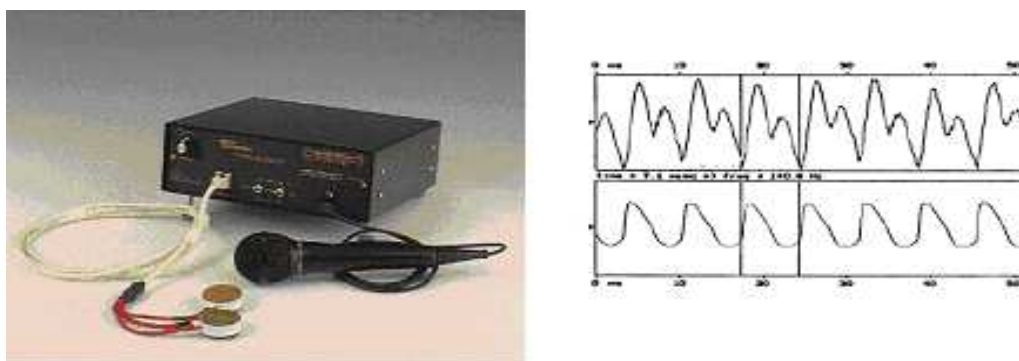


Figura 3.3: Exemplo mostrando o EGG (na esquerda) e o sinal de voz (traçado superior) e o sinal proveniente do EGG (em baixo).

3.2.2 Locais de Recolha e Características dos Informantes

A gravação do *corpus* foi efectuada, com base em inquéritos no terreno, durante 2002-2003 e em seis regiões diferentes (Figura 3.4) desde o Litoral e Interior Norte (Minho e Trás-os-

Montes) até ao Sul (Alentejo e Algarve). No centro foram contemplados, quer o litoral (Beira Litoral), quer o interior (Beira Interior).

Para cada região foram seleccionadas duas localidades. Deste modo foi possível efectuar uma maior cobertura do território continental e em simultâneo recolher informação sobre regiões geralmente referidas como tendo uma pronúncia característica dos sons nasais.

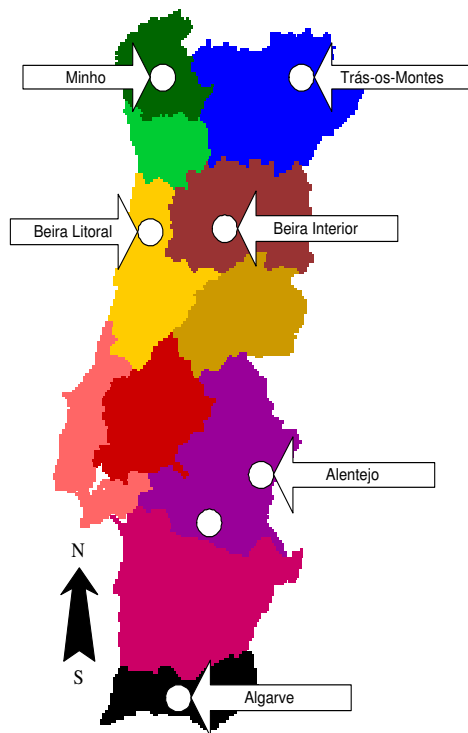


Figura 3.4: Mapa de Portugal com indicação das zonas onde pertence cada um dos informantes.

Em cada região foi recolhida informação de quatro informantes, dois do sexo masculino e dois do sexo feminino. Todos os informantes preenchiam os seguintes requisitos:

- serem naturais e residentes da localidade escolhida, ou, pelo menos, terem iniciado aí a sua escolarização;
- saberem ler, mas não terem uma escolaridade superior ao actual ensino obrigatório ou seu correspondente;
- terem uma idade compreendida entre 35 e 70 anos.

Esta preocupação prende-se com o facto de se pretender obter uma amostragem homogénea, de modo a apenas fazer variar o parâmetro geográfico. Note-se ainda que este perfil de informante deverá garantir uma maior representatividade da respectiva variante geográfica.

No Anexo B apresentam-se mais informações sobre os informantes que colaboraram na recolha do *corpus*.

Recolha

As gravações foram sempre realizadas no terreno, utilizando como equipamento: um CSL 4400 e EGG 6103, ambos da Kay Elemetrics Corp., e um computador portátil. O sinal foi captado através de um microfone Shure SM48 directamente para o disco duro. Durante a recolha, foram utilizados, sempre que possível, estímulos visuais, através de imagens (alguns exemplos são apresentados na Figura 3.5 e no Anexo C) que conduzissem o informante a produzir as palavras pretendidas, evitando, deste modo, a leitura. A cada informante foram solicitadas duas repetições do *corpus*.



Figura 3.5: Figuras utilizadas na gravação do *corpus*. As imagens correspondem às palavras *ponte*, *pintor* e *bandeira*, respectivamente.

3.2.3 Software Utilizado

Para este projecto escolheu-se o sistema *Speech Filing System* (SFS) desenvolvido por Mark Huckvale no University College London². Tanto para a gravação, como para a segmentação, foi utilizado o programa SFSWin, o qual possibilita a gravação simultânea dos dois canais de entrada que se pretendem utilizar: neste caso, o do microfone e o do EGG. O SFSWin permite visualizar todas as formas de onda, em simultâneo ou individualmente (por exemplo, o oscilograma, espectrograma e o sinal que representa os movimentos das cordas vocais na produção do som). Para além disso, permite também operar sobre determinados segmentos seleccionados, ouvir, apagar, anotar, etc. O sistema SFS possibilita a conversão dos ficheiros no formato SFS em ficheiros WAV, de modo a poderem ser processados por outros programas, nomeadamente pelo sistema *EMU Speech Database System* (EMU)³.

3.2.4 Segmentação e etiquetagem

Após a gravação, procedeu-se a uma primeira audição, com o objectivo de seleccionar as palavras que acabaram por constituir os ficheiros a analisar. Para tal, foi efectuada a anotação

²<http://www.phon.ucl.ac.uk/resource/sfs>

³<http://emu.sourceforge.net>

manual do início e do fim de cada item e posteriormente, com o auxílio, do comando SFS **wordchop**, criaram-se, de uma forma automática, os ficheiros individuais.

O material segmentado tem a duração total de 1.78 horas.

A maior parte dos ficheiros individuais foram mais tarde etiquetados, de modo a delimitar o segmento com interesse e o respectivo contexto. Por exemplo, no ficheiro contendo a palavra *tinto* [tĩtu] foi anotada a posição da vogal nasal [ĩ], bem como do [t] anterior e posterior. Esta anotação foi efectuada utilizando o alfabeto fonético SAMPA para Português. Um exemplo anotado da palavra *pintor* é apresentado na Figura 3.6.

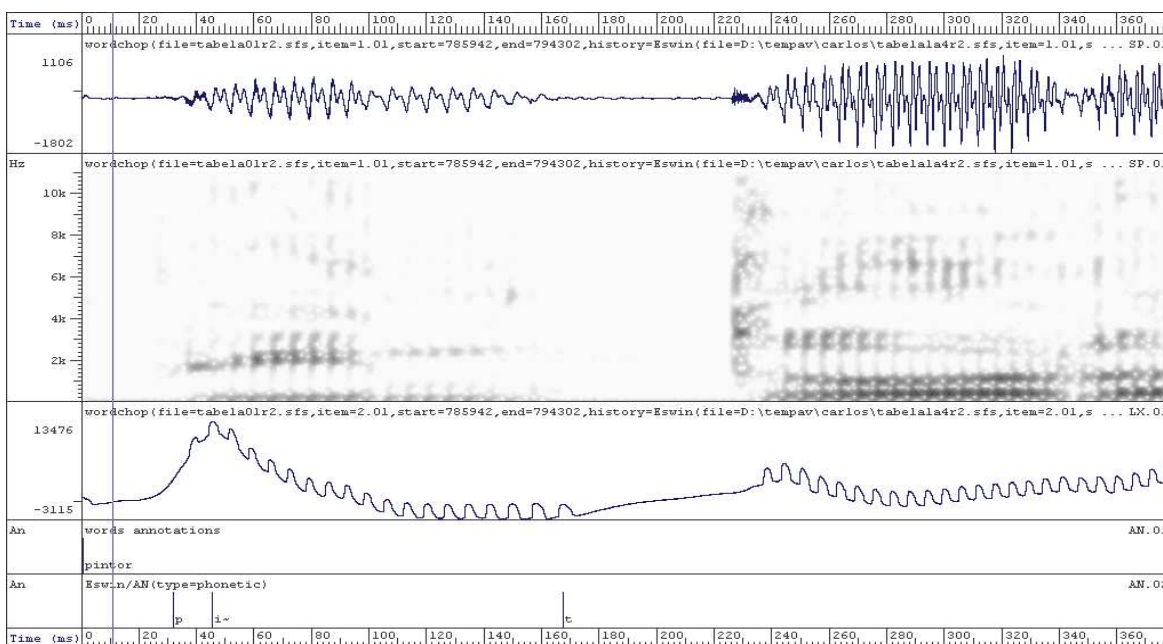


Figura 3.6: Exemplo de uma anotação do *corpus*. De cima para baixo, pode ver-se o sinal de voz, o sinal do EGG, anotação ao nível das palavras e anotação fonética. Nesta última pode ver-se os pontos de início da vogal nasal e da consoante que a segue ([t]) assim como a indicação de qual o segmento que antecede a vogal nasal (neste caso o [p]).

Actualmente, apenas as tabelas 1 a 4 (vogais nasais entre oclusivas) e a tabela 9 (vogais orais) estão anotadas para todos os informantes gravados. Após uma primeira anotação, efectuada com o auxílio do SFS, todos os ficheiros foram verificados por um linguista profissional. Este processo resultou num total de 3000 amostras de vogais nasais e um pouco mais de 1400 vogais orais anotadas.

A Figura 3.7 apresenta a distribuição das amostras anotadas por vogal e região. É notório que a distribuição por região é quase uniforme. Por seu lado, a distribuição por vogal está longe de ser uniforme: a nasal [ĩ] é a mais representada, com um total de 958 amostras, enquanto que o [ũ] apenas tem 303 amostras anotadas no total.

A distribuição por género, não apresentada na figura, favorece ligeiramente os homens com 1615 amostras de oradores masculinos, contra 1405 de mulheres.

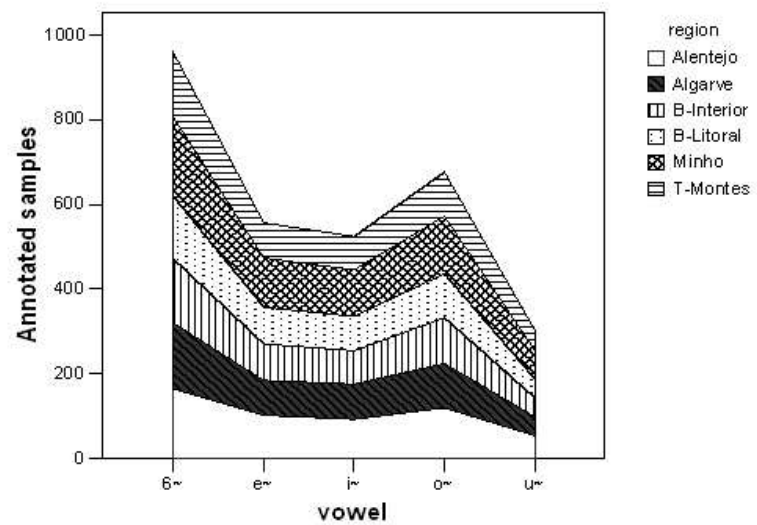


Figura 3.7: Distribuição das amostras anotadas por vogal e por região.

Capítulo 4

Características da Fonte

Para a continuidade do trabalho em síntese articulatória, em desenvolvimento na Universidade de Aveiro, torna-se imprescindível a obtenção de dados sobre a produção dos sons do PE, em particular, sobre a produção das diversas vogais nasais.

O sintetizador articulatório necessita, para além da configuração dos articuladores, fora do âmbito deste trabalho, de informação sobre parâmetros relacionados com a excitação glotal, dos quais se destacam a Frequência Fundamental (F0) e o Quociente de Abertura (OQ).

Geralmente, os *corpora* existentes e as análises publicadas sobre este tema, limitam-se a um número reduzido de locutores, regiões do país e contextos fonéticos. No entanto, o novo *corpus* Regional (apresentado na secção 3.2) contempla todos os contextos fonéticos para as vogais nasais do PE, privilegiando variedades regionais, o que torna possível a obtenção de novos dados significativos.

Análises focando este tipo de parâmetros tinham já sido efectuadas no âmbito do projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português” (A.Teixeira *et al.*, 2004b). No entanto, o estudo foi realizado tendo por base informação relativa a, apenas, 6 dos 25 informantes.

Actualmente, com todo o *corpus* segmentado e anotado torna-se relevante reanalisar estes parâmetros.

Nas secções seguintes (4.1 e 4.2) introduz-se o assunto em análise e apresenta-se o método utilizado no estudo. Os resultados obtidos podem ser encontrados na secção 4.3. Tendo em vista a obtenção de informação sobre os parâmetros de excitação glotal sem a influência do factor informante, foram analisados os valores obtidos após uma normalização. Este trabalho é apresentado na secção 4.4. A discussão dos diversos resultados (4.5), bem como o resumo do assunto, encerram o capítulo.

4.1 Introdução

Esta secção pretende introduzir os conceitos de F0, OQ e *jitter*, que serão analisados nas restantes secções.

A F0 é definida como a taxa à qual uma onda sonora é repetida por unidade de tempo, o que neste caso equivale ao número de vibrações produzidas pelas cordas vocais por segundo.

A obtenção do valor de F0, com alguma precisão, a partir do sinal acústico, pode ser um problema difícil, devido a diversos factores. Exemplo disso são os casos em que a onda de excitação glotal não é um trem periódico perfeito de impulsos.

Existe um grande conjunto de algoritmos para a obtenção do valor de F0 (Rabiner *et al.* (1976); Silva (1995)). Estes podem ser divididos em três categorias principais, de acordo com o tipo de informação do sinal que é utilizada:

- algoritmos que utilizam as propriedades no domínio do tempo;
- algoritmos que utilizam as propriedades no domínio da frequência;
- algoritmos que utilizam tanto as propriedades no domínio do tempo como as propriedades no domínio da frequência.

Os algoritmos que utilizam as características no domínio do tempo operam directamente sobre o sinal, determinando a F0 através de medições de picos ou vales, passagens por zero ou autocorrelação. Neste tipo de algoritmos está incluído o método da autocorrelação, utilizado neste estudo. Quando o valor de F0 é determinado com base nas características espectrais do sinal de voz, os algoritmos servem-se do facto de que se um sinal é periódico no tempo, então o seu espectro será composto por impulsos com a frequência fundamental e os seus harmónicos.

Por último, os algoritmos híbridos aproveitam, quer as propriedades no domínio do tempo, quer no domínio da frequência. Como exemplo, temos a técnica de filtragem inversa simplificada (SIFT).

Se o sistema fonador fosse um mecanismo ideal e perfeitamente estabilizado, não existiria qualquer diferença nos períodos fundamentais, excepto quando um orador alterasse, propositadamente, o *pitch*. No entanto, este sistema não é, de modo algum, um mecanismo perfeito, e todos os ciclos vibratórios dos oradores são de, alguma forma, imperfeitos.

A perturbação da frequência (ou período) - normalmente designado de *jitter* - é uma medida da variação da frequência fundamental (ou período fundamental) ciclo a ciclo .

Existe uma grande variedade de medidas que podem ser utilizadas para obter o valor do *jitter* (Baken e Orlikoff, 2000). Estas são exploradas de seguida.

A amplitude da perturbação da frequência está correlacionada com a média da frequência fundamental e as maiores diferenças no período entre ciclos sucessivos estão associadas a períodos fundamentais maiores (reciprocamente, maiores frequências fundamentais tendem a ter menor perturbação). Logo, uma medida para o *jitter*, não influenciada pela média da

frequência fundamental, passa por gerar uma razão entre alguma forma de perturbação média e o período médio.

É neste pressuposto que a maior parte dos índices de *jitter* se baseiam.

A medida mais simples, obtida com base na ideia anterior, é dada pela perturbação média dividida pela duração média da onda sonora. Quando esta medida é obtida através do período, é denominada de *jitter ratio* e é dada por:

$$JitterRatio = \frac{\frac{1}{N-1} \left[\sum_{i=1}^{N-1} |P_i - P_{i+1}| \right]}{\frac{1}{N} \sum_{i=1}^N P_i} \times 1000 \quad (4.1)$$

onde

- P_i = período do i^o ciclo em ms;
- N = número de períodos na amostra.

Quando o *jitter* é definido em termos da frequência, isto é, quando é definido como a média do valor absoluto das perturbações na frequência fundamental, normalizada pela média do valor desta, é denominado de *Jitter Factor* (JF), ou seja:

$$JitterFactor = \frac{\frac{1}{N-1} \left[\sum_{i=1}^{N-1} |F_i - F_{i+1}| \right]}{\frac{1}{N} \sum_{i=1}^N F_i} \times 100 \quad (4.2)$$

Uma outra abordagem, para a quantificação da perturbação no período, baseada em estatísticas descritivas, é a *Period Variability Index* (PVI). Este índice requer o cálculo de um Coeficiente de Variação (CV) definido como:

$$CV = \frac{\frac{1}{N} \sum_{i=1}^N [(P_i - \bar{P})^2]}{\bar{P}^2} \quad (4.3)$$

O PVI é igual ao CV multiplicado por 1000.

As medidas apresentadas até agora não tomam em consideração o efeito criado por alterações lentas no F0, e tal deve-se ao facto de se basearem em médias de F0.

A *relative average perturbation quotient*, também denominada *frequency perturbation quotient* tenta suavizar esse efeito. Esta medida quantifica a diferença entre um período e o anterior, após corrigir a tendência geral para o aumento dos períodos, ou seja, esta avalia a diferença entre o valor de um período e o que este teria, caso os períodos aumentassem de uma forma perfeitamente regular.

$$RAP = \frac{\frac{1}{N-2} \sum_{i=2}^{N-1} \left| \frac{P_{i-1} + P_i + P_{i+1}}{3} - P_i \right|}{\frac{1}{N} \sum_{i=1}^N P_i} \quad (4.4)$$

A escolha de uma medida para o cálculo do valor do *jitter* depende sempre da precisão requerida e da sua implementação computacional. Neste trabalho foi utilizado o JF.

Diversos eventos do processo de fonação estão relacionados com eventos (segmentos ou picos) correspondentes do sinal EGG ou da sua derivada (DEGG). Por exemplo, o instante de abertura das cordas vocais e o instante do pico positivo no sinal DEGG são coincidentes, assim como o instante de fecho das cordas vocais e o instante do pico negativo do DEGG (Childers, 2000). Assim, o OQ medido da área glotal é equivalente ao OQ medido através do EGG, sendo:

$$\text{OQ} = \text{duração de abertura das cordas vocais} / \text{duração do período glotal}$$

Na secção seguinte é descrito o método utilizado para obter os parâmetros apresentados.

4.2 Método

A obtenção dos parâmetros relacionados com a excitação glotal, apresentados na secção anterior, foi realizada de uma forma automática sobre todos os ficheiros do *corpus* Regional. Para tal foi utilizado o sistema SFS, em particular a linguagem *Speech Measurement Language* (SML), que permite manipular os dados provenientes deste sistema.

De uma forma geral, a análise baseou-se na procura automática, para cada ficheiro do *corpus*, de todos os segmentos anotados como vogal (oral e nasal) e na obtenção dos parâmetros em estudo: F0, OQ e *jitter*. A informação sobre o falante, a região, o género do falante, a tabela, o item, etc., foi também guardada e posteriormente utilizada em análises estatísticas, realizadas com o auxílio do programa SPSS¹.

De seguida, apresenta-se informação mais detalhada sobre a forma como cada um dos parâmetros foi obtido.

4.2.1 Frequência Fundamental

O SFS permite a obtenção do valor de F0, quer através do sinal de voz (comando **fxanal**), quer através do sinal do EGG (comandos **vtx** e **fx**).

O comando **fxanal** estima a F0, a partir do sinal de voz, utilizando o método de autocorrelação. O algoritmo utilizado é semelhante ao apresentado por Secrest e Doddington (1983).

No segundo caso o processo consiste em obter informação sobre o início de cada período do sinal do EGG, usando para tal o comando **vtx**. O valor de F0 é obtido com base nesta informação e através do programa **fx**. O comando **vtx** estima a posição dos períodos de excitação glotal a partir do sinal do EGG com base no método utilizado no *Voiscope*² que

¹<http://www.spss.com>

²Instrumento baseado no laringógrafo que possibilita a visualização da informação de F0 no ecrã de um osciloscópio.

consiste em:

1. filtragem
2. controlo de ganho automático
3. detecção de picos.

Este processo é exemplificado na Figura 4.1 onde se pode ver o sinal do EGG, informação sobre o início e fim de cada período e F0.

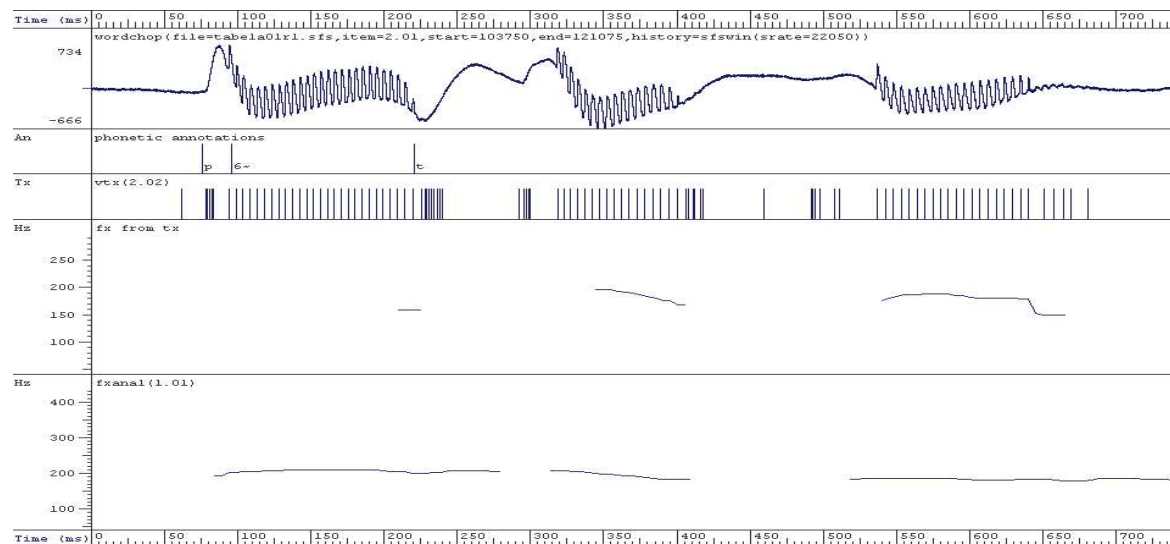


Figura 4.1: Exemplo de cálculo de F0 com base no sinal EGG. Através do sinal do EGG, apresentado no topo, é primeiro calculado o início de cada período glotal (*tx* no sistema SFS) e depois calculada a frequência fundamental (*fx* no sistema SFS). No final é visível a frequência fundamental obtida através do sinal de voz.

Para possibilitar o estudo de uma possível variação deste parâmetro ao longo da realização da vogal, o valor de F0 foi calculado em três pontos: próximo do início (concretamente a 10% da duração da vogal), no meio e próximo do final (a 90% da duração).

4.2.2 Quociente de Abertura

O OQ foi obtido com base em informação sobre a abertura e fecho das cordas vocais gerada pelo comando SFS **txgen**. O processamento efectuado é o seguinte:

1. Filtragem passa-baixo do sinal do EGG
2. Diferenciação do sinal EGG (obtendo DEGG)
3. Procura do máximo e mínimo do sinal DEGG que se encontra acima de um determinado limiar. A posição destes é utilizada como posição de fecho e início de abertura

Com base nesta informação foi desenvolvido um programa em SML para calcular o OQ. Infelizmente a detecção efectuada pelo **txgen** apresenta muitas deficiências, tornando difícil a obtenção de uma boa estimativa de OQ em ficheiros com o sinal EGG de menor qualidade.

4.2.3 *Jitter*

Como se viu anteriormente, existem várias medidas para o cálculo do valor do *jitter* (medida da perturbação do período de oscilação das cordas vocais).

O sistema SFS inclui um comando que permite obter esta informação - **txstat**. No entanto, devido à falta de documentação, não foi possível obter informação sobre o método utilizado por este comando. Optou-se, então, por desenvolver um novo programa de cálculo, utilizando a SML e, devido à sua simplicidade, a medida *Jitter Factor* (Equação 4.2).

O valor do JF é, então, calculado usando uma janela (de comprimento n). De modo a estudar eventuais variações ao longo da vogal, o JF foi calculado:

1. considerando toda duração da vogal;
2. com uma janela de comprimento fixo ($n = 5$) no início;
3. com uma janela de comprimento fixo ($n = 5$) no meio;
4. com uma janela de comprimento fixo ($n = 5$) no fim.

É notório, em algumas realizações das vogais, um decréscimo do valor de F0 ao longo da vogal. Deste modo, o valor do *jitter* terá tendência para atingir valores mais elevados (originados pela utilização de um menor valor da média de F0 na normalização (ver página 31)). Por este motivo, foi também analisado o valor do numerador da fórmula de cálculo do JF.

Os resultados obtidos, para este e outros parâmetros, são apresentados de seguida.

4.3 Resultados

4.3.1 Frequência Fundamental

Após uma primeira análise dos resultados, verificou-se que o processo baseado no EGG não foi capaz de determinar correctamente o valor de F0 para um elevado conjunto de informantes. Este definiu como nulos 37% dos valores de F0 calculados para o início da vogal (F0START), 46% dos valores obtidos no meio da vogal (F0MED) e 33% dos correspondentes ao final da vogal nasal (F0END). Deste modo, os valores de F0 obtidos através deste processo não foram utilizados na análise estatística final. Esta análise foi efectuada com base, apenas, nos resultados do processo baseado no sinal de voz.

Os valores de F0 inferiores a 100 e superiores a 300Hz foram também eliminados, uma vez que estão em análise dados provenientes de informantes masculinos e femininos. Esta

selecção, bem como a análise estatística dos dados, foi efectuada com o auxílio do programa SPSS.

Começou por se estudar a variação de F0 ao longo das vogais (início, meio e fim) em função da vogal, considerando a nasalidade e o acento. Este estudo é apresentado na Figura 4.2. Os gráficos superiores da figura apresentam os resultados obtidos para vogais orais, enquanto que os da esquerda são relativos a vogais não acentuadas.

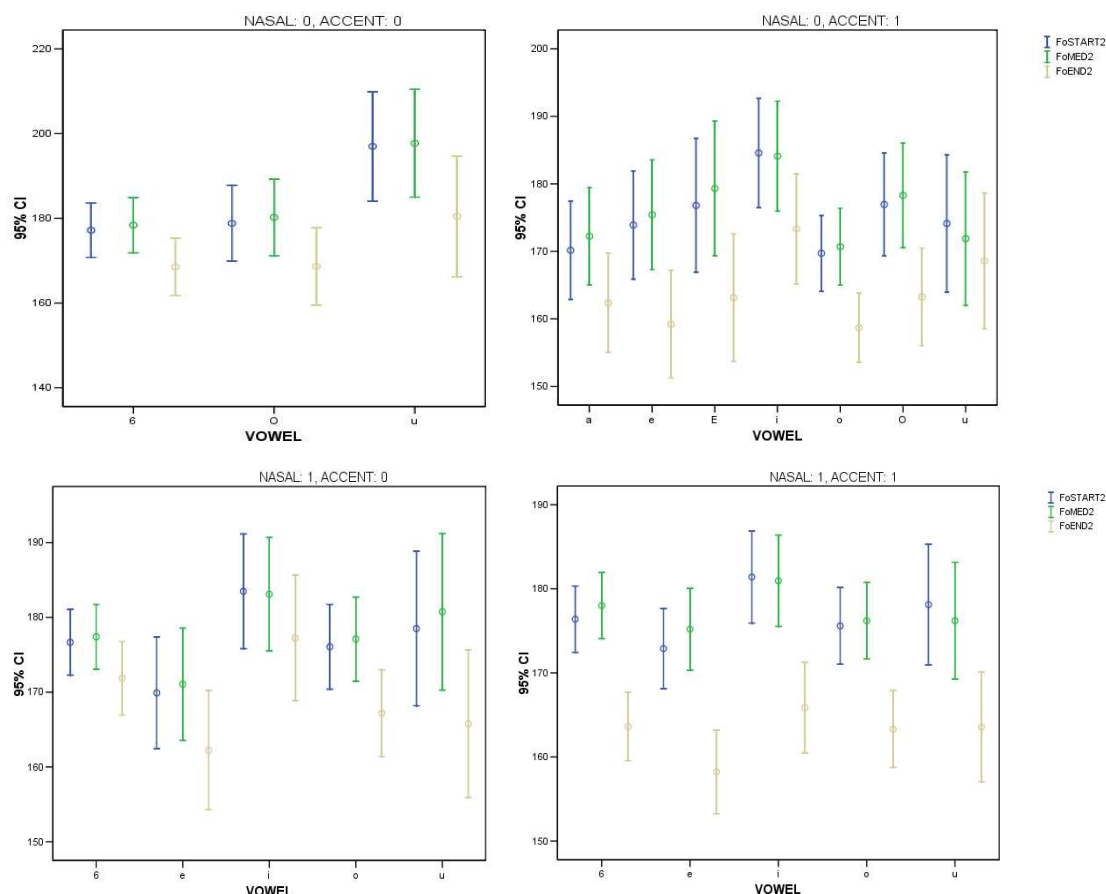


Figura 4.2: Intervalo de Confiança para os valores de F0 (em Hz) em função da vogal, da nasalidade (nasais em baixo), do acento (acentuadas nos dois gráficos da direita).

É possível verificar uma diminuição do valor de F0 no final da realização das vogais, especialmente notório para vogais nasais acentuadas. Esta diminuição é também visível na tabela 4.1, onde se apresentam os valores obtidos nas diferentes fases das vogais analisadas, para cada um dos falantes.

Na tabela 4.2, que apresenta a diferença entre os valores final e inicial para cada vogal, verifica-se que, para todas as vogais, o valor médio final é cerca de 10Hz inferior ao valor médio inicial. Sendo que o valor médio para vogais orais é de -10,35Hz e -11,82Hz para as nasais.

A Figura 4.3 apresenta o estudo da variação de F0 para os vários informantes. O estudo

Orador	F0 Inicial	F0 Medial	F0 Final
jqm	158,24	160,98	144,64
jsd	157,04	154,78	141,04
ars	245,73	247,95	224,66
alt	223,96	224,37	217,53
amt	174,64	176,59	150,42
jlm	131,32	132,54	122,81
apm	135,42	136,34	120,32
nmt	152,54	154,24	139,49
acr	207,69	208,19	195,28
jlz	131,36	131,12	115,62
ccd	213,70	211,33	202,48
jrd	112,05	111,30	111,50
ral	212,23	212,79	206,30
clr	144,22	144,32	132,17
prq	123,75	125,40	119,32
ocn	200,78	201,32	201,32
nsb	185,44	187,07	178,78
mmc	210,62	214,23	193,28
adl	202,60	201,73	192,02
ald	137,44	139,78	123,62
als	126,60	126,55	121,19
apc	114,08	114,15	114,63
dro	122,09	120,87	117,17
cva	247,84	245,80	228,78
scf	178,23	181,19	173,63

Tabela 4.1: Valores médios de F0 (Hz) medidos no início, meio e fim da vogal para cada um dos 25 informantes.

	Diferença
a	-6
6	-8,57
e	-13,71
E	-12,1
i	-12,46
o	-10,92
O	-12,74
u	-13,18

Tabela 4.2: Diferença entre os valores inicial e final de F0 (Hz).

foi efectuado para as várias vogais nasais, tendo em conta a acentuação.

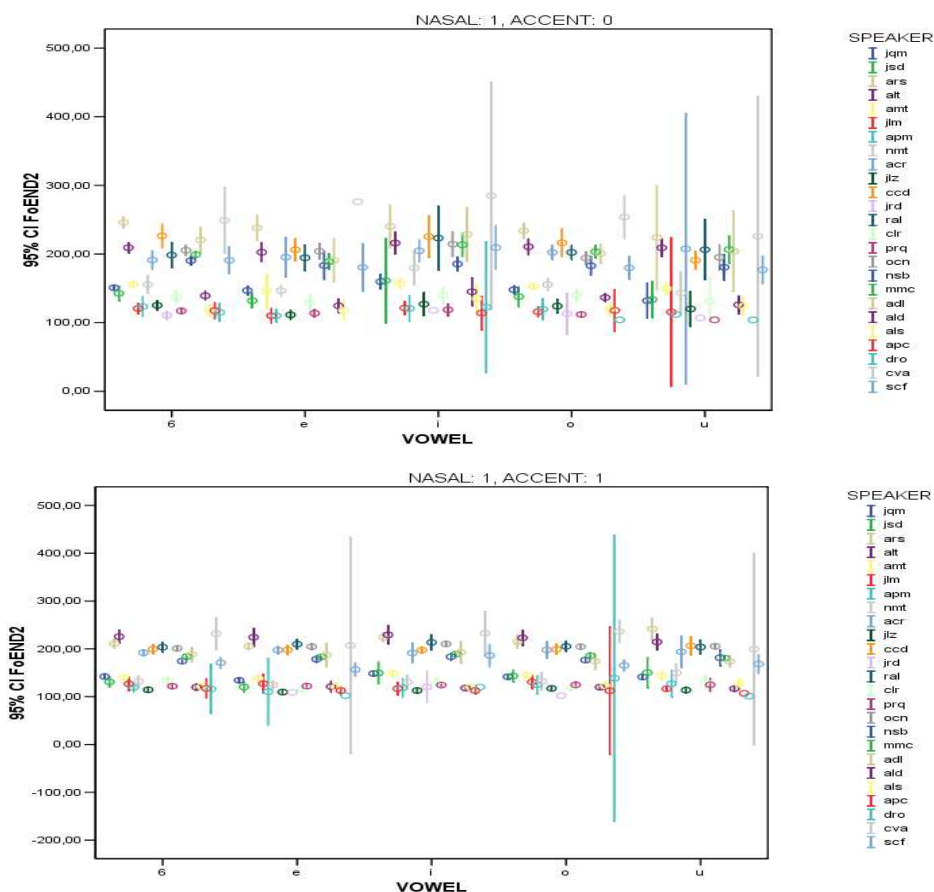


Figura 4.3: Intervalo de Confiança para os valores de F0 (Hz) das vogais nasais (acentuadas ou não) em função da vogal e do falante.

Neste estudo nota-se uma elevada dispersão nos valores médios e intervalos de confiança para a mesma vogal. Não consideramos que exista um valor médio por vogal, pois existe sobreposição dos factores intrínsecos ao falante. Esta influência significativa do factor informante é confirmada por testes estatísticos de análise de variância (ANOVA).

4.3.2 Quociente de Abertura

A figura 4.4 apresenta o diagrama de extremos e quartis (*boxplot*) dos valores do OQ em função da vogal e da nasalidade.

É possível notar uma reduzida variação, quer para os valores medianos, quer para a dispersão dos valores. Isto é válido, quer para as vogais orais, quer para as correspondentes nasais.

Na figura 4.5 é apresentado o *boxplot* dos valores do OQ para os vários informantes, os quais estão sumariados nas tabelas 4.3 e 4.4. A análise do gráfico indica uma elevada variação dos valores médios, bem como uma grande dispersão para alguns, poucos, informantes. Sobressai,

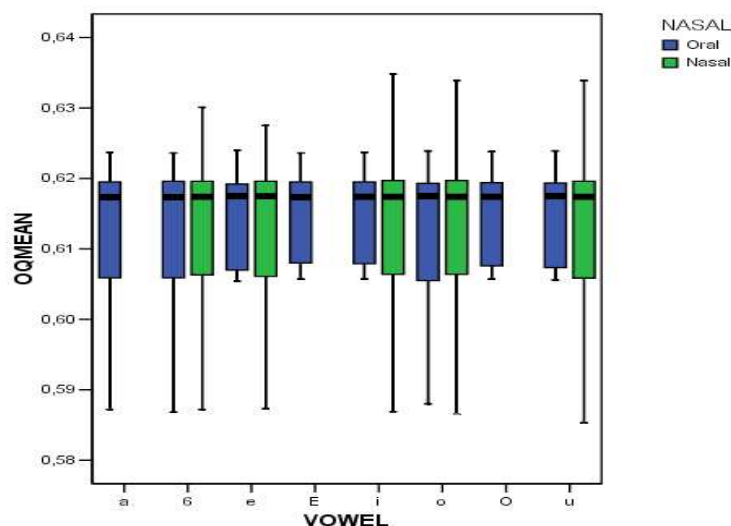


Figura 4.4: Diagrama de extremos e quartis (*boxplot*) dos valores do quociente de abertura em função da vogal e nasalidade.

ainda, o valor inferior do OQ para os falantes ADL, ALD e ALS da região do Minho, bem como para o falante ALT do Alentejo. Esta diferença é notória, quer para as vogais orais, quer para as nasais.

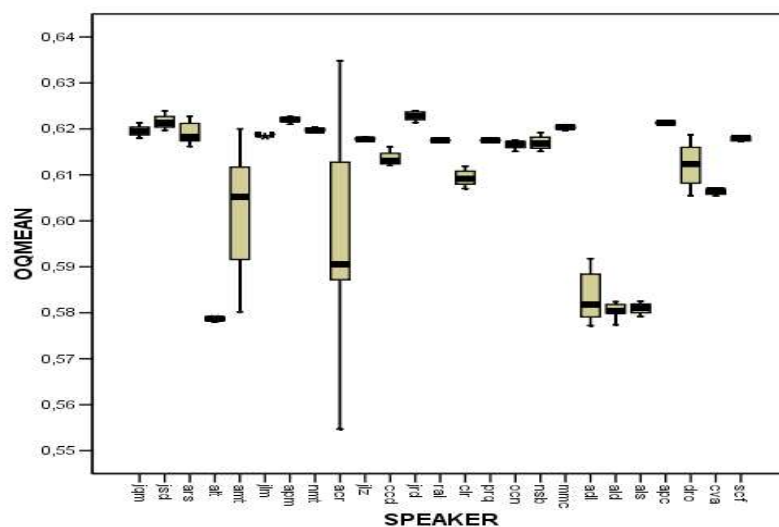


Figura 4.5: *Boxplots* do quociente de abertura para cada falante.

4.3.3 Jitter

Não foi possível obter os valores do JF para alguns informantes. O problema ficou a dever-se à dificuldade que o comando SFS **hqtx** (utilizado no programa SML desenvolvido) tem em processar sinais EGG de menor qualidade. Deste modo, os valores apresentados de seguida

	a	6	e	E	i	o	O	u
jqm	0,62	0,58	0,62	0,62	0,62	0,62	0,62	0,62
jsd	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
ars	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
alt	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58
amt	0,61	0,61	0,62	0,61	0,61	0,62	0,61	0,62
jlm	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
apm	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
nmt	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
acr	0,59	0,59	0,58	0,59	0,59	0,58	0,59	0,58
jlz	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
ccd	0,61	0,61	0,61	0,61	0,61	0,61	0,61	0,61
jrd	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
ral	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
clr	0,61	0,61	0,61	0,61	0,61	0,61	0,61	0,61
prq	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
ocn	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
nsb	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
mmc	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
adl	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58
ald	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58
als	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58
apc	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
dro	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62
cva	0,60	0,61	0,61	0,61	0,61	0,61	0,61	0,61
scf	0,62	0,62	0,62	0,62	0,62	0,62	0,62	0,62

Tabela 4.3: Valores Médios de OQ para as vogais orais e por falante. Na primeira coluna apresenta-se a sigla pela qual é identificado o falante.

	6	e	i	o	u
jqm	0,62	0,62	0,62	0,62	0,62
jsd	0,62	0,62	0,62	0,62	0,62
ars	0,62	0,62	0,62	0,62	0,62
alt	0,58	0,58	0,58	0,58	0,58
amt	0,59	0,59	0,60	0,60	0,60
jlm	0,62	0,62	0,62	0,62	0,62
apm	0,62	0,62	0,62	0,62	0,62
nmt	0,62	0,62	0,62	0,62	0,62
acr	0,60	0,61	0,61	0,61	0,60
jlz	0,62	0,62	0,62	0,62	0,62
ccd	0,61	0,61	0,61	0,61	0,61
jrd	0,62	0,62	0,62	0,62	0,62
ral	0,62	0,62	0,62	0,62	0,62
clr	0,61	0,61	0,61	0,61	0,61
prq	0,62	0,62	0,62	0,62	0,62
ocn	0,62	0,62	0,62	0,62	0,62
nsb	0,62	0,62	0,62	0,62	0,62
mmc	0,62	0,62	0,62	0,62	0,62
adl	0,58	0,59	0,58	0,59	0,58
ald	0,58	0,58	0,58	0,58	0,58
als	0,58	0,58	0,58	0,58	0,58
apc	0,62	0,62	0,62	0,62	0,62
dro	0,61	0,61	0,61	0,61	0,62
cva	0,61	0,61	0,61	0,61	0,61
scf	0,62	0,62	0,62	0,62	0,62

Tabela 4.4: Valores Médios de OQ para vogais nasais e por falante.

devem ser analisados tendo em conta este facto.

Nas Figuras 4.6, 4.7 e 4.8 apresentam-se os diagramas de extremos e quartis dos valores do JF obtidos para o início, meio e fim da vogal, respectivamente, e em função da vogal e da nasalidade.

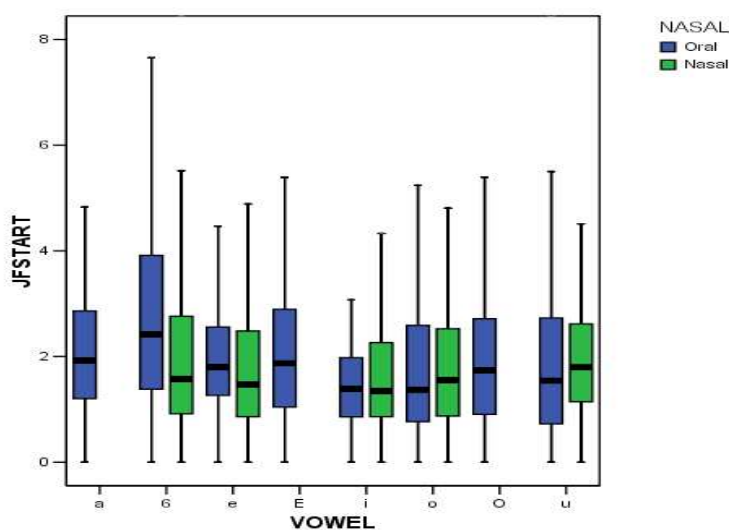


Figura 4.6: *Boxplots* do valor do Jitter Factor na posição inicial da vogal. São apresentados valores para cada vogal separados pela nasalidade.

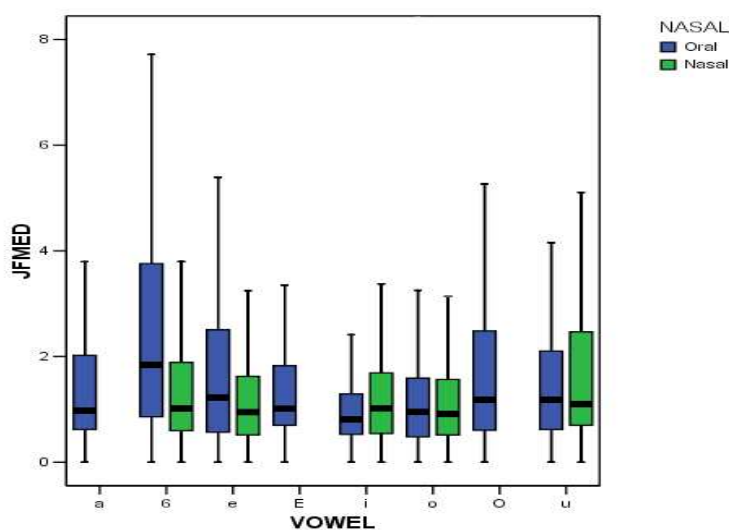


Figura 4.7: *Boxplots* do valor do Jitter Factor na posição medial da vogal. São apresentados valores para cada vogal separados pela nasalidade.

Os valores obtidos para a totalidade da vogal, em função da vogal e da nasalidade são apresentados na Figura 4.9.

Nesta primeira análise gráfica, verifica-se uma reduzida variação dos valores obtidos. No entanto, após uma análise mais detalhada dos valores médios correspondentes, apresentados

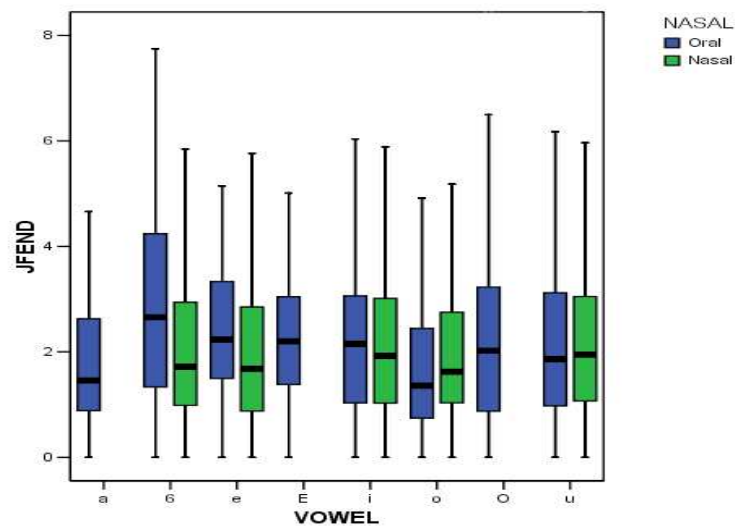


Figura 4.8: *Boxplots* do valor do Jitter Factor na posição final da vogal. São apresentados valores para cada vogal separados pela nasalidade.

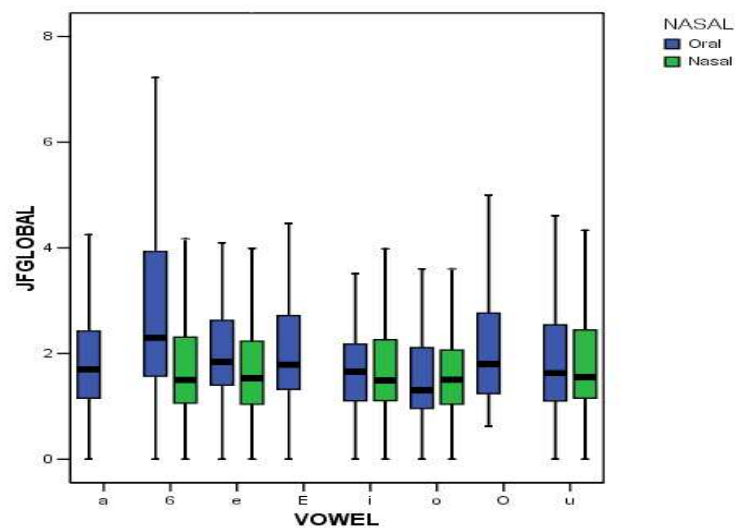


Figura 4.9: *Boxplots* do valor do Jitter Factor usando a totalidade da duração da vogal. São apresentados valores para cada vogal separados pela nasalidade.

na tabela 4.5, evidencia-se a existência de um valor inferior do JF medido a meio da vogal, relativamente ao início e fim, sendo o valor final o mais elevado.

Estes valores contemplam a análise por vogal tendo em conta a nasalidade, nas três posições medidas (JFSTART, JFMED, JFEND) e na globalidade da vogal (JFGLOBAL).

	a	6	e	E	i	o	O	u
oral	2,16	2,87	2,16	2,30	1,61	1,95	1,96	1,99
	1,40	2,50	1,70	1,56	1,09	1,33	1,72	1,66
	1,87	2,98	2,42	2,28	2,39	1,83	2,18	2,28
	2,05	2,85	2,15	2,06	1,78	1,74	2,14	2,08
nasal	,	2,05	1,83	,	1,85	1,87	,	2,12
	,	1,53	1,33	,	1,37	1,33	,	1,78
	,	2,19	2,07	,	2,15	2,04	,	2,28
	,	1,90	1,81	,	1,84	1,76	,	2,02

	Média
JFSTART	1,9930595
JFMED	1,5005002
JFEND	2,1697886
JFGLOBAL	1,92428114

	Médias	
	Oral	Nasal
JFSTART	2,1200406	1,9419101
JFMED	1,6392915	1,4444812
JFEND	2,2557149	2,1352677
JFGLOBAL	2,10506703	1,85222505

Tabela 4.5: Valores médios obtidos para o Jitter Factor. Na primeira tabela por vogal e tendo em conta a nasalidade. Na tabela intermédia médias globais para os valores medidos nas três posições e na globalidade da vogal. Na tabela de baixo à esquerda os valores para vogais orais, à direita para vogais nasais.

No entanto, estando o valor do JF dependente do valor médio de F0, é necessário determinar se essa diminuição se deve ao aumento deste último. Os valores médios do numerador da expressão de cálculo do JF são apresentados na Figura 4.10. Estes continuam a indicar um menor valor para medida obtida a meio da vogal.

De uma forma geral, os resultados confirmam os obtidos em fases anteriores do projecto (A.Teixeira *et al.*, 2004b).

4.4 Análise após normalização

Dos resultados apresentados anteriormente, não é possível retirar um valor médio de cada parâmetro por vogal, pois parece existir uma influência significativa do informante nos valores obtidos. Isto é especialmente notório na análise efectuada aos valores de F0.

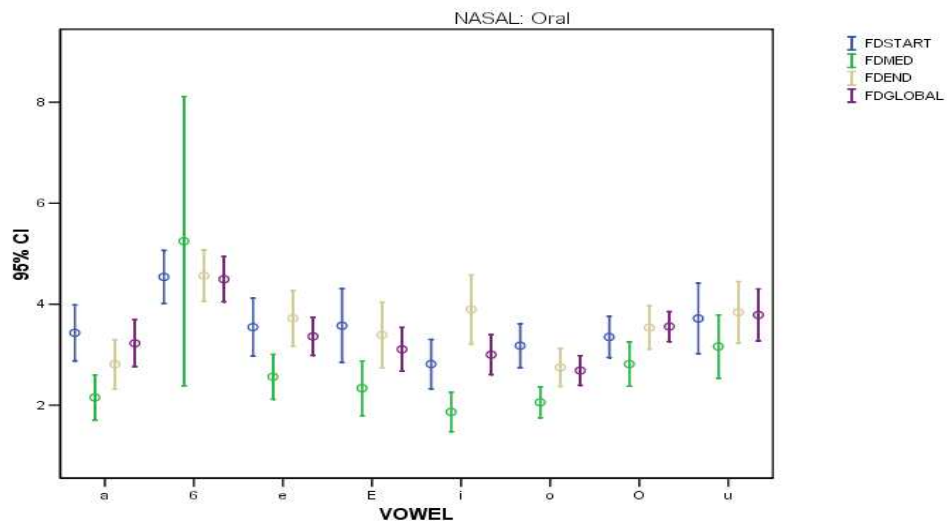


Figura 4.10: Numerador do cálculo de JF para orais.

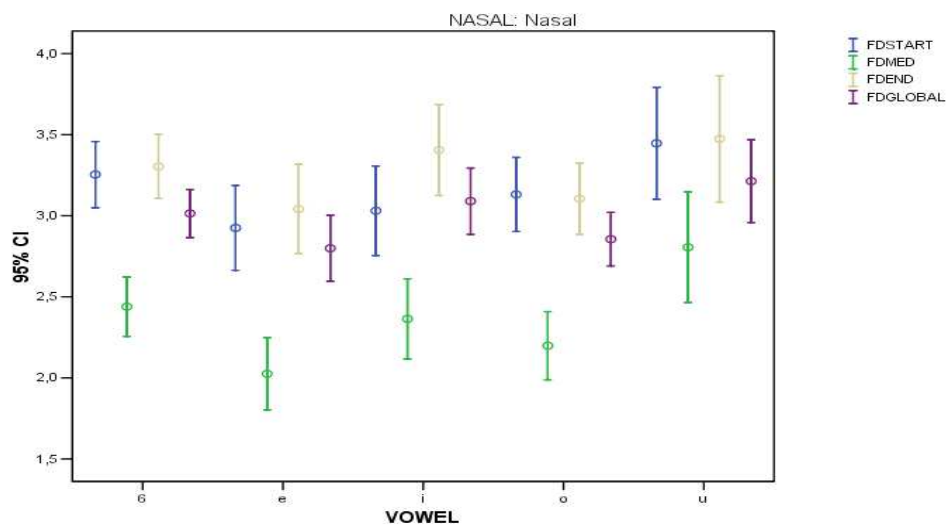


Figura 4.11: Numerador do cálculo de JF para nasais.

Deste modo, foi desenvolvido um estudo cujo objectivo passa por analisar os valores dos parâmetros glotais após uma “diminuição” do factor informante.

É este estudo que preenche as secções seguintes.

4.4.1 Método

De modo a retirar uma eventual influência do falante nos resultados obtidos, foram desenvolvidos estudos contemplando a mesma informação após uma normalização pelo valor mediano de cada falante.

Esta normalização foi inicialmente efectuada com o SPSS. No entanto, como a obtenção destes valores é mais simples através do programa de análise estatística R³, optou-se pela utilização deste, nesta fase do trabalho.

A fraca qualidade dos ficheiros com o sinal proveniente do EGG impossibilitou a obtenção do valor do JF para alguns dos informantes. Devido a este facto a análise apresentada nesta secção não pôde ser estendida a este parâmetro, tendo incidido apenas sobre os valores de F0 e OQ.

Devido à elevada quantidade de resultados existentes, são apresentados nesta secção, apenas, os considerados mais relevantes, encontrando-se os restantes no Anexo D.

4.4.2 Resultados

Frequência Fundamental

A Figura 4.12 apresenta o diagrama de extremos e quartis para o valor de F0, obtido na parte medial da vogal nasal (F0MED), em função do falante. De acordo com aquilo que se obteve anteriormente (figura 4.3), existe uma elevada dispersão nos valores apresentados, o que pode indicar uma influência de factores intrínsecos ao falante.

A normalização dos valores (Figura 4.13) indica a existência de uma ligeira diferença no valor de F0MED para as vogais nasais [ĩ] e [ũ], sendo esta diferença particularmente notória no caso da vogal nasal [ĩ].

Os resultados obtidos para os valores de F0 obtidos nas restantes fases da vogal nasal são apresentados no Anexo D.1.

Quociente de Abertura

O diagrama de extremos e quartis para o valor do OQ, em função dos informantes, foi já apresentado na figura 4.5, demonstrando a elevada variação dos valores medianos deste parâmetro. O gráfico 4.14 apresenta os valores normalizados, para cada vogal nasal, não se notando neste uma grande variação, mas sim uma elevada dispersão.

³<http://www.r-project.org>

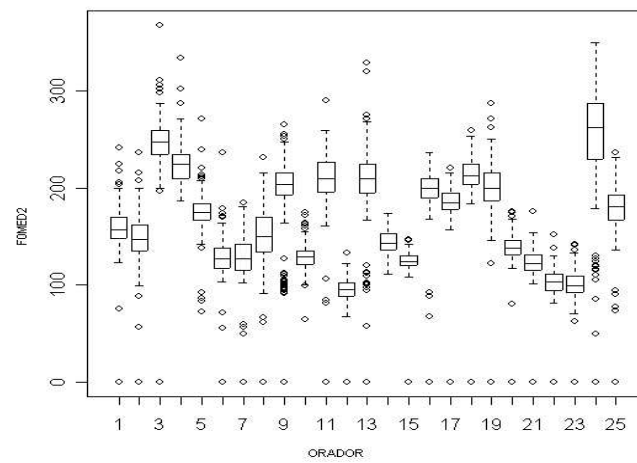


Figura 4.12: Diagrama de extremos e quartis dos valores de F0MED em função do falante.

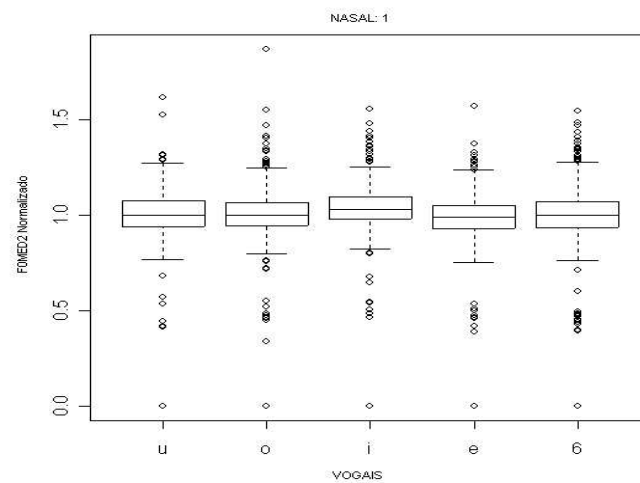


Figura 4.13: Diagrama de extremos e quartis dos valores de F0MED em função das vogais nasais após divisão pela mediana de F0MED.

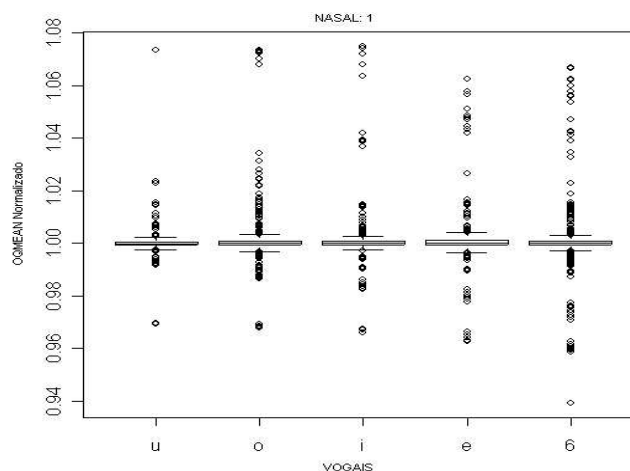


Figura 4.14: Diagrama de extremos e quartis dos valores de OQMean em função das vogais nasais após divisão pela mediana de OQMean.

4.5 Discussão

Os resultados apresentados foram obtidos com uma amostra significativa quanto à variação regional, com a vantagem adicional de se basearem num *corpus* recolhido através de inquéritos no terreno, integrando, por isso, mais variação do que o usual.

Na leitura dos resultados é necessário ter em consideração o facto de estes terem sido obtidos para todas as vogais nasais existentes, nos vários contextos em que estas existem. No entanto, também é de notar que apenas se encontram anotadas, para a totalidade dos informantes gravados, as tabelas 1 a 4 (vogais nasais entre oclusivas) e a tabela 9 (vogais orais).

A existência de informação sobre o comportamento das cordas vocais, obtida através do sinal do EGG, foi relevante para a obtenção de parâmetros como o OQ e o *jitter*. Contudo, a fraca qualidade de alguns ficheiros do EGG impediu a obtenção dos referidos parâmetros, para alguns informantes, em particular o *jitter*.

A comparação entre os resultados obtidos neste estudo e os apresentados em fases anteriores do projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português” (A.Teixeira *et al.*, 2004b) é a única passível de ser efectuada. A utilização de metodologias distintas da actual, e análise de parâmetros diferentes dos estudados neste trabalho, impedem a comparação com outras investigações publicadas.

4.6 Resumo

A existência de informação sobre a frequência fundamental, o quociente de abertura e o *jitter* é essencial para a síntese articulatória, em particular, para a produção, através desta tecnologia, das diversas vogais nasais.

O estudo apresentado neste capítulo teve por base dados recolhidos especialmente para o efeito, reunidos num *corpus*, denominado, ao longo deste documento de, Regional. A existência neste *corpus* de gravações do sinal proveniente do EGG, bem como a recolha de fala oriunda das diferentes regiões de Portugal Continental, permitiu a obtenção de informação relevante, pouco contemplada em estudos anteriores.

A análise desenvolvida baseou-se no processamento dos ficheiros que constituem o *corpus* através do sistema SFS e na, posterior, análise estatística dos resultados com os programas SPSS e R. No que diz respeito aos resultados obtidos, estes indicam a existência de um valor de F0 inferior no final da realização das vogais, não podendo, no entanto, ser definido um valor médio deste parâmetro por informante. Os valores do OQ indicam, de uma forma geral, que cada informante mantém um valor estável à variação da vogal, e por último, a análise do *jitter* indica que o valor do JF medido a meio da vogal é inferior ao obtido para os restantes instantes desta.

Capítulo 5

Caracterização Temporal

Após a análise de factores relacionados com a excitação glotal na produção de vogais nasais, efectuada no capítulo anterior, torna-se também relevante analisar um dos parâmetros mais robustos na identificação deste tipo de sons - a *duração*.

Estudos anteriores indicam que a duração das vogais nasais portuguesas é superior à das correspondentes orais (por exemplo, os de de Sousa (1994) e Delgado-Martins (2002)). No entanto, estes estudos baseiam-se em dados relativos a um número reduzido de informantes, não considerando as variedades regionais existentes no país. Assim, a existência de novos dados reunidos e tratados, quer no *corpus* Regional, quer no *corpus* EMMA, justifica a reanálise deste parâmetro.

Os resultados obtidos em análises anteriores (por exemplo, Teixeira e Vaz (2001); Silva (1995) e Drenska (1988)) apontam para que as vogais nasais se realizem em três diferentes fases, sendo estas: uma vogal oral, seguida de uma breve fase de transição, em que a nasalidade se sobrepõe à vogal, terminando num murmúrio nasal.

Em particular, para vogais nasais entre oclusivas orais, o velo inicia fechado, abre durante a realização vogal e inicia o movimento de fecho necessário para a produção oclusiva seguinte. Os articuladores orais começam o movimento de abertura no início da vogal, e após estarem abertos durante parte da vogal nasal, fecham próximo do fim, para a produção da oclusiva seguinte.

É a duração destes eventos de abertura e fecho, dos articuladores orais e do velo, que será estudada na segunda parte deste capítulo (secção 5.2.1).

A obtenção destes últimos parâmetros só é geralmente possível graças à utilização de tecnologias como a EMMA, técnica, no entanto, inviável para um número alargado de falantes. Assim, é de todo o interesse a obtenção de informação, sobre as fases referidas anteriormente, com base apenas no sinal acústico. Este estudo exploratório finaliza o capítulo.

5.1 Duração Total

5.1.1 Introdução

No que diz respeito à duração total das vogais nasais, é interessante notar que estudos realizados sobre línguas que contêm vogais nasais fonologicamente distintas, apontam para uma maior duração das vogais nasais em relação às orais (ver pag. 12).

Estudos recentes efectuados para o português (de Sousa, 1994) demonstram que as vogais nasais são sempre mais longas que as orais, e que a “sílabas” (a medida do primeiro pulso glotal da consoante precedente ao último período regular da vogal em questão) em que as nasais se encontram são igualmente mais longas que a “sílabas” em que se localizam as vogais orais. Esta maior duração reflecte-se, de maneira mais ténue, também ao nível da palavra.

As análises efectuadas numa fase preliminar do *corpus* Regional, em que apenas estavam recolhidos e segmentados dados relativos a 6 informantes, confirmaram como significativa a influência da vogal, nasalidade, acento e posição na duração. No entanto, quer a influência da região, quer a influência da repetição, não se revelaram significativas.

Nas secções seguintes é apresentado o processo desenvolvido para a obtenção da duração total das vogais nasais recolhidas, seguindo-se a apresentação e discussão dos resultados obtidos (5.1.2 e 5.1.3).

5.1.2 *Corpus* Regional

Embora a análise do parâmetro *duração* tivesse já sido efectuada no âmbito do projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português” (A.Teixeira *et al.*, 2004b), o aumento do número de informantes gravados, bem como do número de dados segmentados e anotados, justifica a reanálise deste parâmetro.

O método utilizado na obtenção de tais valores é apresentado de seguida.

Método

Como a anotação efectuada aos ficheiros do *corpus* indica onde começa a vogal e onde começa o segmento que a segue, o processo de obtenção da duração é muito simples, envolvendo apenas a utilização do comando **lengthn** do SFS, que retorna o comprimento do segmento pretendido.

Foi, assim, efectuada uma procura automática em todos os ficheiros do *corpus* Regional, dos segmentos anotados como vogal (oral e nasal) e para cada segmento foi guardada informação, quer sobre a sua duração, quer sobre os vários factores intervenientes (falante, região, género, tabela, repetição, etc.). Estes dados foram, posteriormente, utilizados nas análises estatísticas, realizadas com o programa SPSS e cujos resultados são apresentados de seguida.

Resultados

A primeira análise com interesse passa pela obtenção dos valores médios das durações das vogais nasais. Na tabela 5.1 apresentam-se os valores médios e os desvios padrão para as 5 vogais nasais, utilizando os 25 falantes, das 6 regiões. Os intervalos de confiança, a 95%, dos valores médios em análise estão representados na figura 5.1.

	[6̃]	[ẽ]	[ĩ]	[õ]	[ũ]	<i>p</i>
\bar{d}_t	160,54	180,24	173,52	167,00	167,15	0.000
S_{d_t}	43,28	42,47	43,25	47,05	54,31	

Tabela 5.1: Duração Total (ms).

A tabela 5.1 mostra que existe uma influência significativa da vogal nasal na duração ($p < 0.05$), isto é, d_t é dependente da vogal. Desta tabela é ainda possível verificar que as vogais [ẽ] e [ĩ] apresentam as maiores durações, enquanto que a menor duração é a da nasal [6̃]. É de ainda de notar a semelhança nos valores médios das durações de [õ] e [ũ].

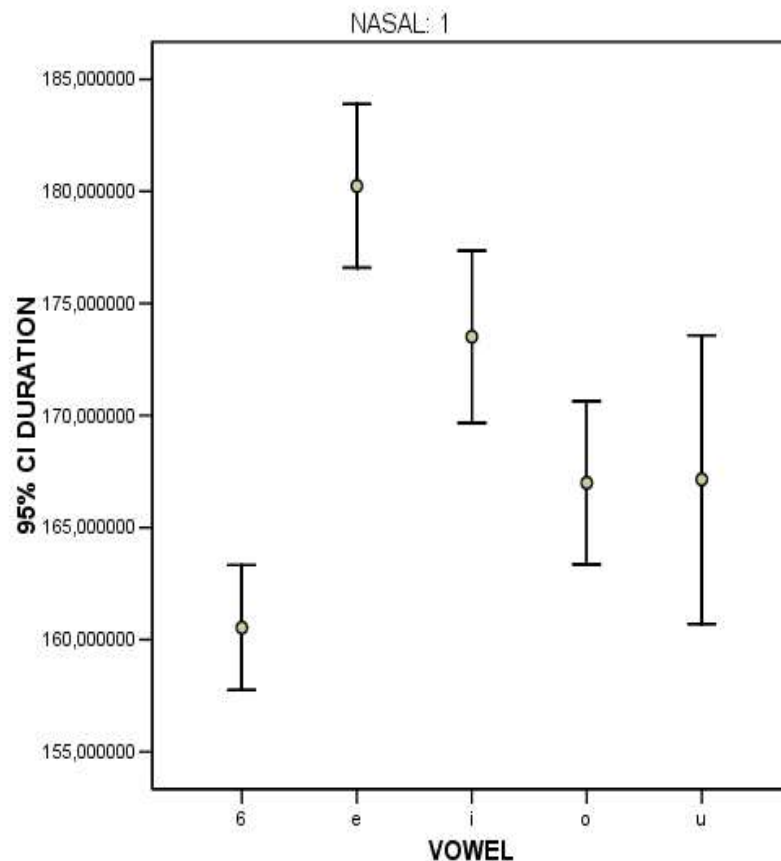


Figura 5.1: Intervalo de confiança a 95% dos valores médios da duração (ms) das vogais nasais do *corpus* Regional.

Os valores médios obtidos, para as oito vogais, em função da nasalidade e acentuação são

apresentados na tabela 5.2.

	Oral		Nasal	
	não acentuada	acentuada	não acentuada	acentuada
a	,	122,85	,	,
6	72,67	,	138,06	177,04
e	,	120,03	156,06	188,82
E	,	118,09	,	,
i	,	115,55	145,19	185,50
o	,	131,80	140,10	184,40
O	88,38	128,52	,	,
u	72,22	127,09	122,52	185,36

Tabela 5.2: Durações médias das vogais analisadas (ms).

A dispersão dos valores médios das durações, apresentada sobre a forma de intervalo de confiança, é apresentada na Figura 5.2, onde é possível verificar que as vogais orais não acentuadas apresentam a menor duração média e que as vogais nasais em posição acentuada as maiores durações.

A dispersão dos valores médios das durações para as vogais nasais, tendo em consideração as várias regiões, é apresentada na Figura 5.3. O gráfico indica uma menor duração para os falantes da região do Algarve, em particular para vogais nasais não acentuadas, enquanto que os falantes de Trás-os-Montes apresentam, em geral, durações superiores.

Este resultado é também visível na Tabela 5.3, onde estão apresentadas as médias para cada vogal nasal em função da região e da acentuação.

	Alentejo	Algarve	Beira Interior	Beira Litoral	Minho	Trás-os-Montes
6 não acentuada	140,33	136,44	131,99	140,11	132,15	147,81
acentuada	171,17	154,70	191,71	183,35	178,17	186,79
e não acentuada	148,75	153,14	147,24	155,18	149,82	184,83
acentuada	172,60	164,16	201,94	194,42	192,85	210,98
i não acentuada	137,86	140,54	148,16	141,21	137,90	165,55
acentuada	174,36	168,95	198,10	187,44	183,94	204,03
o não acentuada	139,20	135,57	150,66	142,40	130,72	143,50
acentuada	180,27	169,42	192,64	190,97	180,00	195,13
u não acentuada	135,86	127,83	123,41	117,52	104,77	127,64
acentuada	179,18	159,54	211,77	180,42	185,34	199,14

Tabela 5.3: Médias de duração para as vogais nasais das 6 regiões, a negrito encontram-se os valores superiores (ms).

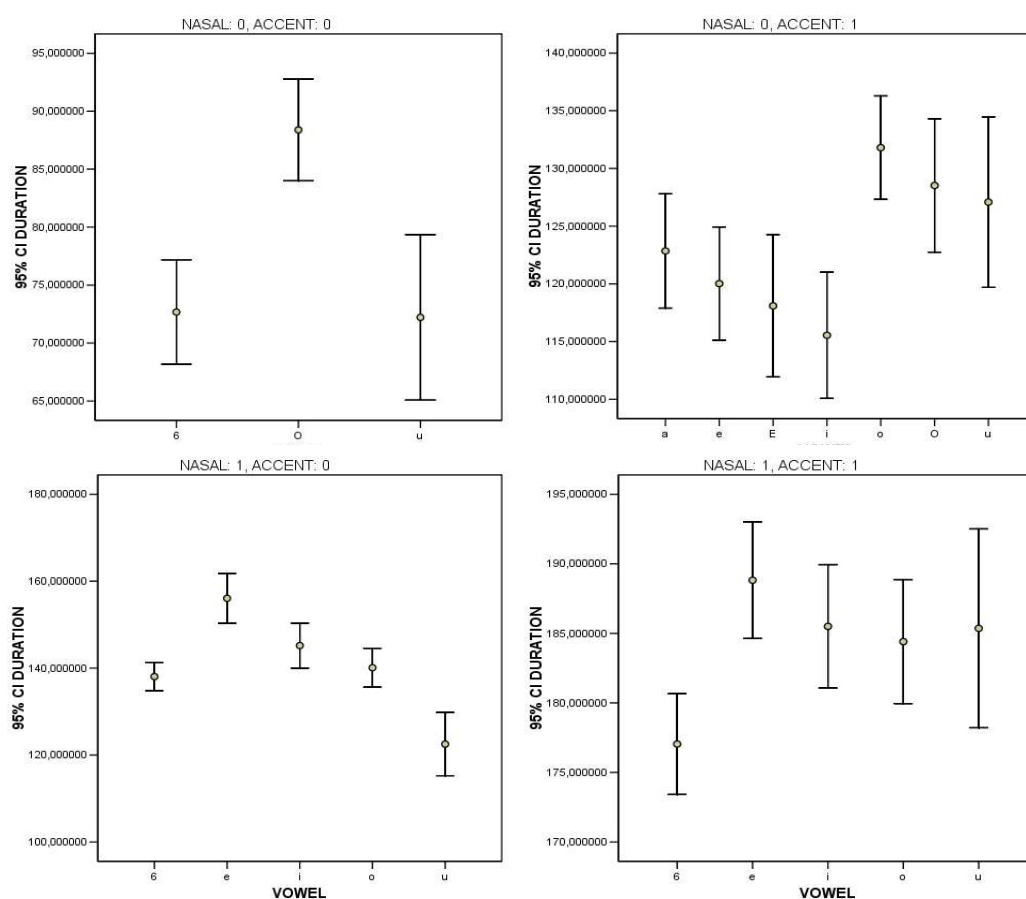


Figura 5.2: Intervalo de Confiança a 95% e valor médio da duração das vogais orais e nasais (ms). Os dois gráficos de cima referem-se a vogais orais (nasal=0), os dois gráficos da esquerda referem-se ao caso não acentuado (accent=0).

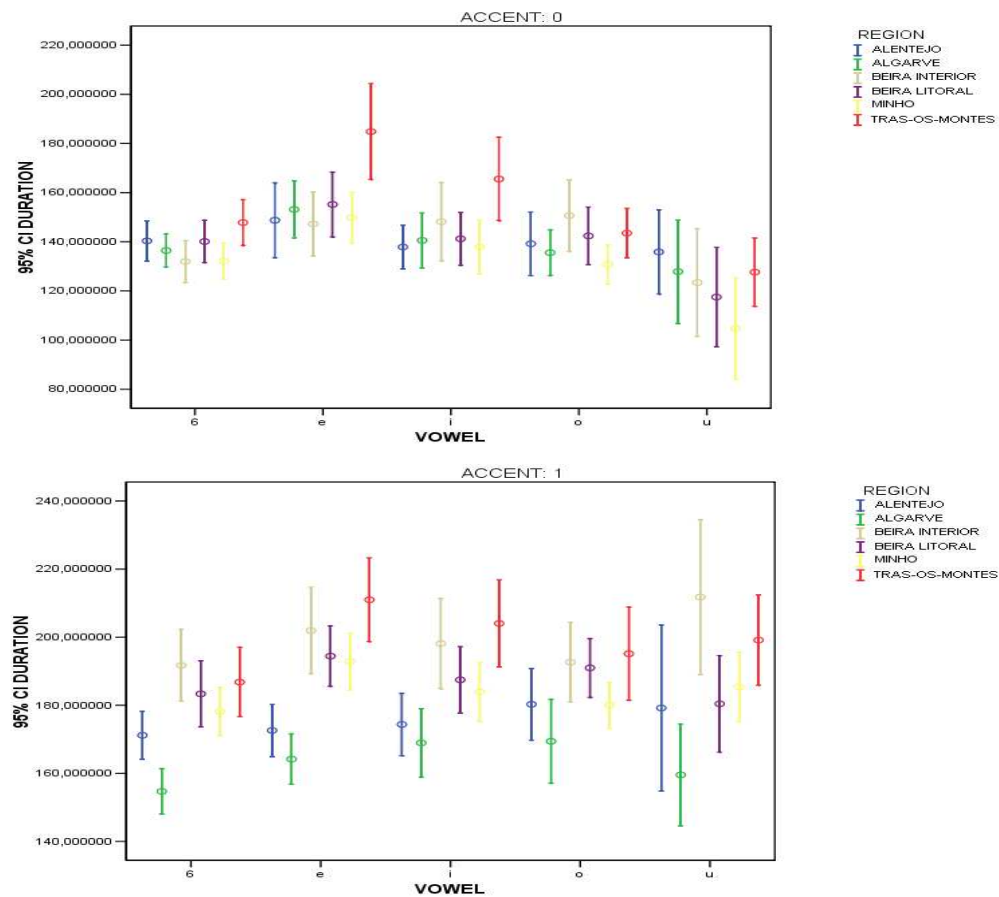


Figura 5.3: Intervalo de Confiança das durações (ms) para as vogais nasais em posição não acentuada (em cima) e acentuada (em baixo) apresentando separadamente os valores obtidos para cada região.

5.1.3 *Corpus* EMMA

A análise apresentada de seguida continua o estudo relativo ao parâmetro *duração*, incidindo agora sobre as vogais nasais existentes no *corpus* EMMA.

Este *corpus* foi gravado apenas para um informante, o que faz com que a única análise com interesse diga respeito à influência da vogal na duração.

Método

A obtenção da duração foi efectuada com o auxílio do sistema SFS, em particular da linguagem SML, referida anteriormente. O valor da duração foi obtido através da função `sml lengthn(pattern, n)` que retorna o comprimento (em segundos) do *n*-ésimo segmento, anotado conforme indicado.

A análise estatística dos valores obtidos é apresentada de seguida.

Resultados

A única análise efectuada consiste na obtenção dos valores médios e desvios padrão das durações de cada uma das vogais nasais existentes no *corpus* em análise. Estes valores estão sistematizados na Tabela 5.4. À semelhança do que aconteceu no estudo efectuado sobre o *corpus* Regional, verifica-se a influência significativa da vogal nasal na duração ($p < 0.05$), isto é, d_t é dependente da vogal.

	[ɐ̃]	[ẽ]	[ĩ]	[õ]	[ũ]	<i>p</i>
\bar{d}_t	271,2	266,6	238,3	277,8	249,4	0.015
S_{d_t}	43,63	26,45	41,70	38,67	29,80	

Tabela 5.4: Duração Total (ms).

Os intervalos de confiança, a 95%, dos valores médios das durações são apresentados graficamente na figura 5.4. A figura permite concluir que os maior e menor valores para as durações, foram encontrados, respectivamente, para as nasais [õ] e [ĩ]. Estes resultados demonstram uma elevada disparidade para amplitude dos valores obtidos nesta análise e na anterior.

A primeira hipótese levantada para justificar tal diferença, foi a ligação do informante do *corpus* EMMA às regiões da Beira Litoral e do Minho. Assim, as figuras 5.5 e 5.6 apresentam os intervalos de confiança a 95% para os valores médios da duração dos informantes masculinos das referidas regiões, do *corpus* Regional. Com esta análise pretende-se associar as características do informante do *corpus* EMMA, com as características dos informantes das regiões referidas do *corpus* Regional. No entanto, a análise das figuras não permite associar qualquer semelhança entre os valores apresentados nestas e os obtidos para o informante do *corpus* EMMA. Mais uma vez, quer a amplitude dos valores obtidos, quer os valores máximo

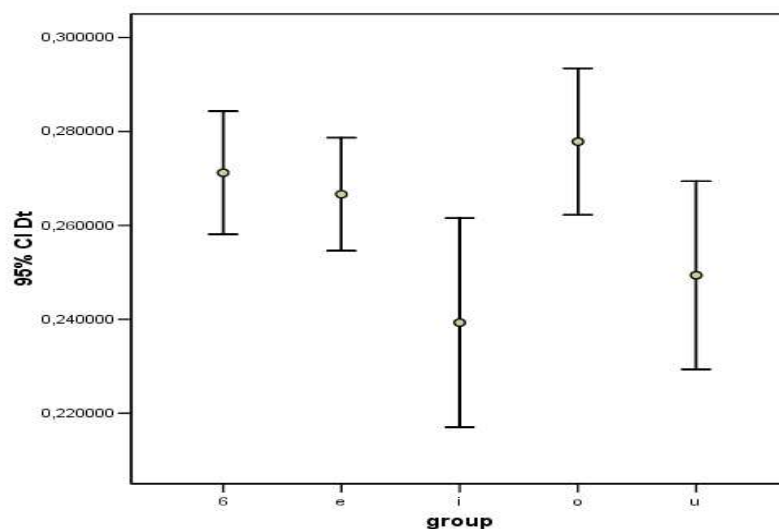


Figura 5.4: Intervalo de confiança a 95% para os valores médios das durações das vogais nasais do *corpus* EMMA (s).

e mínimo, não se assemelham aos obtidos com o *corpus* EMMA.

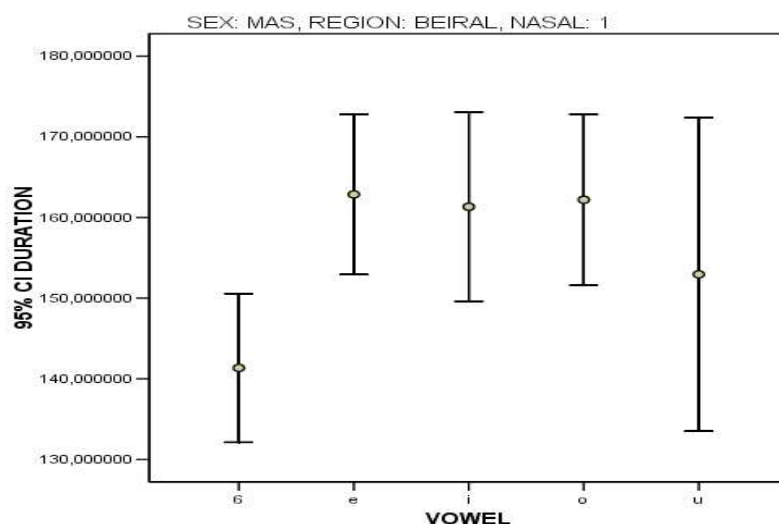


Figura 5.5: Intervalo de confiança a 95% para os valores médios das durações das vogais nasais dos informantes masculinos, da região da Beira Litoral, do *corpus* Regional (ms).

5.1.4 Discussão

Os resultados anteriores apresentam valores para as durações totais das vogais nasais do PE totalmente distintos entre si. Na análise efectuada com os dados provenientes do *corpus* Regional obtiveram-se valores médios para as durações nunca superiores a 190 *ms*, enquanto que os valores obtidos com o *corpus* EMMA não são nunca inferiores a 230 *ms*. As diferenças

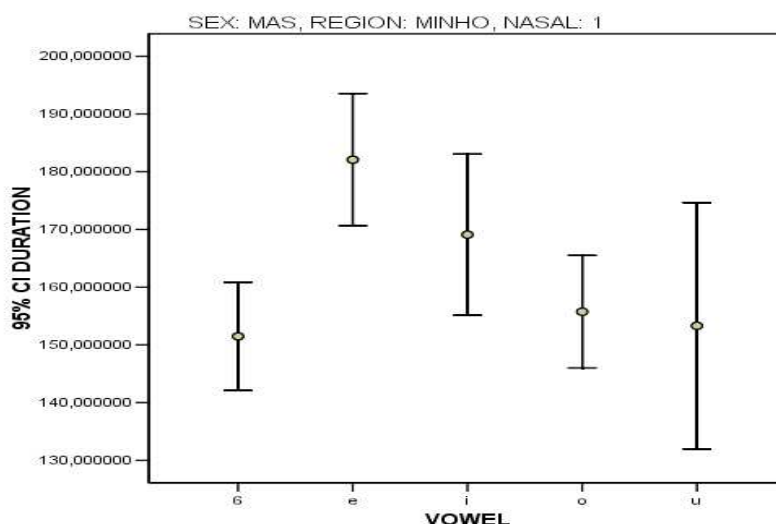


Figura 5.6: Intervalo de confiança a 95% para os valores médios das durações das vogais nasais dos informantes masculinos, da região do Minho, do *corpus* Regional (ms).

reflectem-se também ao nível do valor máximo e mínimo obtido. No caso do *corpus* Regional são as vogais nasais [ɐ] e [e] que apresentam, respectivamente, o menor e maior valores obtidos. Nos dados provenientes do *corpus* EMMA o menor e o maior valores da duração total são obtidos para as vogais [i] e [o].

A ligação do informante do *corpus* EMMA às regiões da Beira Litoral e Minho foi também analisada. No entanto, quer a amplitude dos valores obtidos para os informantes do *corpus* Regional das regiões referidas, quer mesmo a análise dos maior e menor valores obtidos, não permitem justificar a diferença existente. Esta disparidade, pode ser justificada, no entanto, pelo facto de o *corpus* EMMA ter sido gravado apenas com um falante, e nas condições inerentes a qualquer recolha através deste sistema.

Ambas as análises apresentam como significativa a influência da vogal nos valores das durações totais das vogais nasais.

5.2 Obtenção de informação acerca das Fases de uma vogal nasal

A caracterização temporal das vogais nasais também passa pela obtenção de informação detalhada sobre as diferentes fases que as constituem.

As vogais nasais em contexto $C\tilde{V}C$ iniciam com uma configuração onde domina a oralidade e terminam com uma configuração nasal dominante (Teixeira e Vaz, 2001). Deste modo, estamos interessados em obter a duração da parte inicial da vogal em que o véu palatino se mantém fechado ou quase fechado, bem como a duração do período de abertura e fecho deste e a duração da parte final entre o fecho dos articuladores orais e o fecho do velo.

Esta análise será efectuada na secção 5.2.1. A secção 5.2.2 apresenta um estudo exploratório realizado com o objectivo de obter informação semelhante à referida acima, através da avaliação da correlação entre as zonas das vogais nasais com características espectrais estáveis e os eventos anotados nos sensores que controlaram o velo e a passagem oral no sistema EMMA.

5.2.1 Utilizando anotações dos sensores EMMA

Foi referido anteriormente que para vogais nasais entre consoantes orais, o velo inicia fechado, abre, e durante a produção da vogal pára e inicia o movimento de fecho necessário para a produção da consoante que lhe segue.

Por seu lado, a passagem oral inicia a produção da vogal nasal aberta, e após estar nessa posição durante parte da vogal nasal, fecha perto do fim, de modo a produzir a consoante oral seguinte. Estes eventos estão representados na Figura 5.7

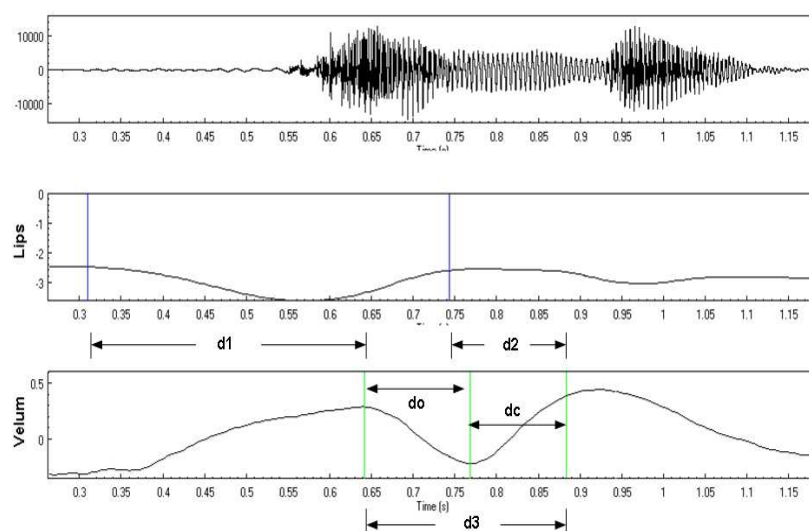


Figura 5.7: Eventos do velo e orais para uma vogal nasal entre consoantes orais.

O que se pretende analisar nesta secção é a duração dos intervalos associados a cada um destes eventos. O método utilizado é descrito de seguida.

Método

Foi referido anteriormente (secção 3.1) que, de modo a facilitar a análise posterior, os dados do *corpus* EMMA foram anotados usando quatro níveis: palavra, fonemas, eventos relativos ao velo e eventos relativos à passagem oral.

As anotações relativas à palavra e aos fonemas tinham já sido introduzidas em todos os ficheiros. No entanto, as anotações referentes aos eventos do velo e eventos orais estavam incompletas. Deste modo, foi necessário proceder à anotação de todos os ficheiros contendo

palavras com vogais nasais entre oclusivas orais (os ficheiros que tinham já sido anotados foram reanotados de modo a uniformizar o método aplicado). A anotação foi efectuada recorrendo às facilidades existentes no sistema EMU, originando um total de 41 ficheiros anotados, contendo 164 palavras.

A análise desenvolvida baseia-se no apresentado em (Teixeira e Vaz, 2001) e foi realizada, mais uma vez, com o auxílio do sistema SFS, com o qual foi criado um ficheiro de dados contendo a duração dos vários intervalos associados aos eventos de abertura e fecho dos articuladores orais e velo, isto é:

- o tempo entre a abertura oral e o início do movimento de abertura do velo - d_1
- a duração do intervalo existente entre o fecho dos lábios e o fecho do velo - d_2
- a duração do movimento de abertura do velo - d_0
- a duração do movimento de fecho do velo - d_c
- a duração do ciclo abertura-fecho do velo - d_3

A duração total da vogal nasal - d_t - foi já analisada anteriormente (secção 5.1.3).

Prevenindo alguma variação dos valores das durações referidas, devida à duração total da vogal nasal, foi também calculado o valor das durações normalizadas segundo d_t .

Os resultados são apresentados de seguida.

Resultados

As tabelas seguintes resumem os resultados obtidos. A tabela 5.5 apresenta os valores médios obtidos para as durações em análise, não normalizadas. Cada um dos valores apresentados representa a distância do tipo especificado na linha, calculada para a vogal em análise na respectiva coluna. A última coluna apresenta o nível de significância obtido para o teste ANOVA (Guimarães e Cabral, 1997), tendo como factor a vogal.

	[6̃]	[ẽ]	[ĩ]	[õ]	[ũ]	p
d_1	90,7	99,6	64,5	70,1	39,3	0,064 (ns)
d_2	112,9	105,3	72,9	91,0	114,2	0,000 (***)
d_o	181,4	173,1	158,4	176,5	163,1	0,261 (ns)
d_c	125,8	128,9	106,3	117,5	128,8	0,223 (ns)
d_3	307,2	302,0	264,7	293,9	291,9	0,056 (ns)

Tabela 5.5: Durações Fases (ms).

Na tabela 5.5 é possível observar que a duração d_2 , isto é, a duração do intervalo entre o fecho oral e o fecho do velo, é influenciada pela vogal nasal em análise. Por seu lado, o valor da duração do ciclo abertura-fecho do velo, d_3 , apresenta um valor de p muito próximo

de 0,05, sendo, no entanto, superior a este valor, o que torna a dependência da vogal não significativa.

As durações normalizadas, segundo d_t , de cada uma das fases das vogais nasais são apresentadas na tabela 5.6.

	[6̃]	[ẽ]	[ĩ]	[õ]	[ũ]	p
d_{1n}	33,3	37,9	28,6	26,4	17,0	0,118 (ns)
d_{2n}	41,8	39,8	31,1	32,3	46,5	0,000 (***)
d_{on}	67,7	64,6	66,5	63,2	65,9	0,658 (ns)
d_{cn}	46,8	48,9	45,0	42,2	52,1	0,177 (ns)
d_{3n}	114,5	113,5	111,5	105,4	118,0	0,102 (ns)

Tabela 5.6: Durações Fases (%).

Mais uma vez se verifica, quer a influência significativa da vogal em d_2 , mesmo após a eliminação do factor d_t , isto é, do factor duração total, quer a não significância do factor vogal para os restantes factores.

Os valores médios para a razão, r , entre d_0 e d_t , para cada uma das vogais nasais, são apresentados na tabela 5.7. Para este parâmetro, a análise ANOVA, para o factor vogal, apresenta um nível de significância, p , superior a 0,05, o que torna a dependência da vogal não significativa.

	[6̃]	[ẽ]	[ĩ]	[õ]	[ũ]	p
r	1,5	1,6	1,5	1,5	1,3	0,564 (ns)

Tabela 5.7: *Ratio* r entre d_0 e d_t .

Discussão

Na tabela 5.5 verifica-se que apenas o parâmetro d_2 (duração do intervalo entre o fecho dos lábios e o fecho do velo) é influenciado pela vogal. Este resultado mantém-se após a normalização do parâmetro - d_{2n} apresentada na tabela 5.6, o que reforça a percepção inicial de o intervalo entre o fecho oral e o fecho do velo ser dependente da vogal. Os níveis de significância apresentados nas tabelas referidas revelam, ainda, que a dependência dos restantes parâmetros da vogal analisada é não significativa.

Os resultados obtidos em (Teixeira e Vaz, 2001) apresentam algumas diferenças em relação aos valores obtidos neste trabalho, nomeadamente, no que diz respeito aos valores obtidos para as durações normalizadas. Os valores apresentados neste trabalho, são, de uma forma geral, superiores aos apresentados na comunicação referida. Estas diferenças podem, no entanto, ser justificadas pela análise, actual, de um maior conjunto de dados.

5.2.2 Com base apenas no sinal acústico: estudo exploratório

Existindo interesse em obter dados similares aos anteriores para um número alargado de pessoas, foi efectuado um estudo exploratório relacionado com a obtenção de informação, sobre as fases referidas anteriormente, com base apenas no sinal acústico. Este estudo exploratório é justificado pela dificuldade existente em obter um conjunto alargado de dados com base na EMMA, originada pelos vários impedimentos inerentes à tecnologia, como, por exemplo, a colocação do sensor do velo (figura 3.1).

O estudo pretende avaliar o índice de correlação existente entre zonas do sinal de voz com características espectrais estáveis e os eventos assinalados pelos sensores EMMA.

O método utilizado é apresentado de seguida.

Método

Para determinar as zonas do sinal de voz com características espectrais estáveis, foi utilizado o comando SFS **nsegment**. Esta rotina usa um procedimento baseado em programação dinâmica que determina a localização de N-1 limites de um item, de modo que se cada região determinada por tais limites fosse substituída pela sua média, o erro quadrático de distorção seria mínimo. O algoritmo é semelhante ao apresentado por Bridle e Sedgwick (1977).

Esta técnica foi aplicada aos sinais existentes, mais concretamente à parte relativa à vogal nasal dos sinais, de modo a obter-se 3 e 4 segmentos anotados de uma forma automática. Estes números justificam-se pela existência, nas vogais nasais em contexto $C\tilde{V}C$, de uma configuração num plano temporal, em que a parte inicial da vogal nasal pode ser considerada oral, a que se segue a abertura da passagem velo-faríngea e o fecho dos articuladores orais (figura 5.7).

Na fase seguinte foi efectuado um estudo da correlação entre os limites dos segmentos obtidos e os eventos anotados no sensor do velo e nos sensores que controlam a passagem oral (sensores dos lábios e sensores da língua, em particular o colocado na parte posterior). Esta última análise foi efectuada através do programa SPSS.

Devido à extensão dos resultados que foram obtidos, apenas serão apresentados na secção seguinte os mais relevantes, sendo os restantes listados no Anexo E.

A apresentação dos resultados mais significativos, em particular os relativos aos pontos de segmentação com maior interesse, está dividida em duas partes, sendo a primeira referente aos resultados obtidos a partir de uma segmentação em três zonas, enquanto que a segunda parte apresenta os resultados obtidos após a segmentação automática em quatro zonas estáveis.

Resultados após segmentação em três zonas

Os resultados obtidos, reunidos nesta secção, são, em primeiro lugar, apresentados de uma forma gráfica, seguindo-se a apresentação dos valores correspondentes ao índice de correlação

e ao índice de ajustamento linear obtidos. A recta de regressão obtida para cada conjunto de dados é também apresentada.

A figura 5.8 começa por apresentar a relação existente entre o evento de abertura nasal (denominado de NASALOPEN) e o primeiro ponto de segmentação obtido de uma forma automática (SEG1). O elevado índice de ajustamento linear que esta figura parece indicar é confirmado na tabela 5.8. Esta última sumaria o modelo obtido, apresentando o coeficiente de correlação (R) existente entre os dois parâmetros em estudo (NASALOPEN e SEG1) e o coeficiente de ajustamento linear (R square).

A recta de regressão:

$$NasalOpen = 1,011 \times SEG1 - 0,31 \quad (5.1)$$

é mais uma indicadora do elevado ajustamento existente entre estes eventos.

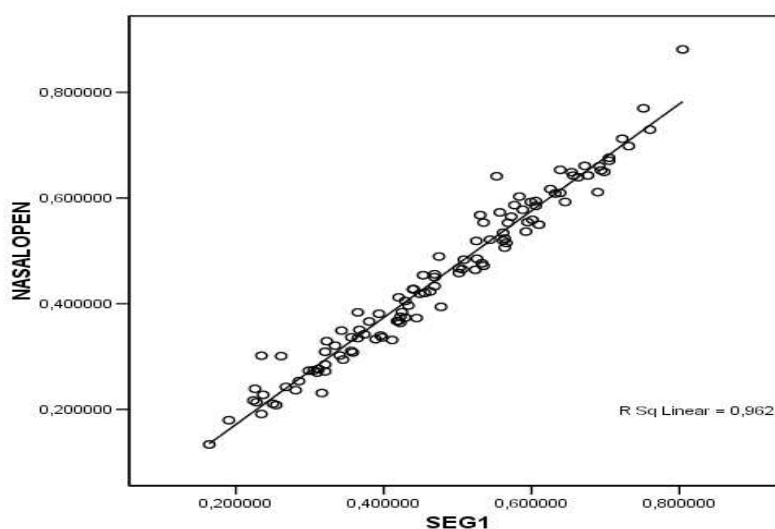


Figura 5.8: *Scatter Plot* entre o evento de abertura do velo (s) e o primeiro segmento automático (s).

Modelo	R	R Square	Adjusted R Square
1	,981	,962	,962

Tabela 5.8: Sumário do Modelo obtido para o evento de abertura do velo e o primeiro segmento automático.

A mesma sequência é apresentada para a relação existente entre o evento de fecho oral (ORALCLOSE) e o segundo segmento obtido de uma forma automática (SEG2). Assim, o gráfico 5.9 apresenta a dispersão entre os dois eventos e tabela 5.8 sumaria o modelo obtido.

Embora a correlação continue a ser elevada, existe uma diminuição no índice de ajusta-

mento linear. A recta de regressão obtida é dada por:

$$OralClose = 1,054 \times SEG2 - 0,23 \quad (5.2)$$

Estes valores continuam a revelar uma elevada dependência linear entre os dois parâmetros em análise.

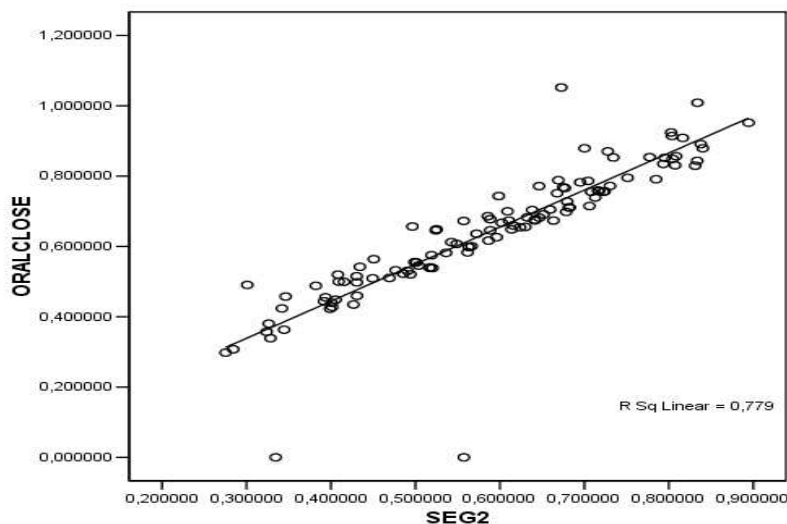


Figura 5.9: *Scatter Plot* entre o evento de fecho dos articuladores orais (s) e o segundo segmento automático (s).

Modelo	R	R Square	Adjusted R Square
1	,882	,779	,777

Tabela 5.9: Sumário do Modelo obtido para o evento de fecho dos articuladores orais (s) e o segundo segmento automático (s).

Resultados após segmentação em quatro zonas

A figura 5.10 apresenta a relação existente entre o evento de abertura nasal (NASALOPEN) e o primeiro segmento obtido de uma forma automática (neste caso o primeiro de quatro, sendo por isso denominado de SEG1.4). Note-se, mais uma vez, a linearidade existente. A tabela 5.10 sumaria o modelo obtido, apresentando o coeficiente de correlação existente entre estes dois parâmetros (R) e o coeficiente de ajustamento linear (R square). A recta de regressão obtida neste caso é:

$$NasalOpen = 1,007 \times SEG1.4 - 0,029 \quad (5.3)$$

Tal como indicava a figura anterior, ambos os conjuntos de valores mencionados corroboram a ideia visual de um elevado índice de correlação e ajuste linear, sendo estes muito semelhantes aos obtidos para a segmentação em 3 zonas efectuada anteriormente.

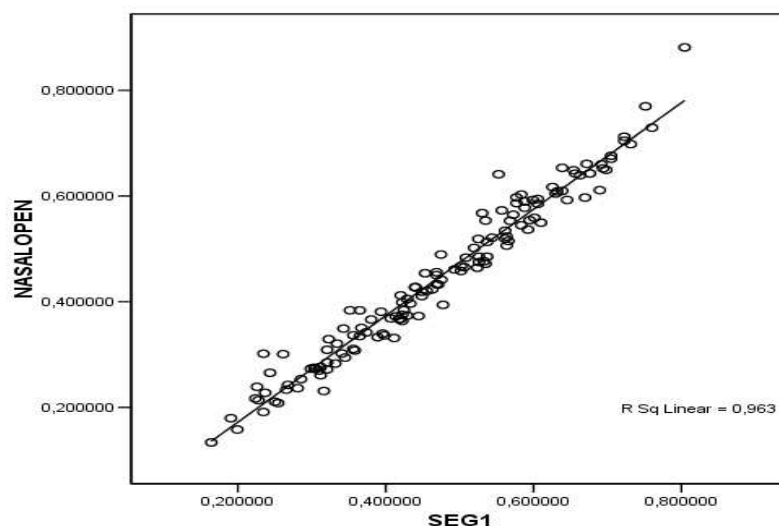


Figura 5.10: *Scatter Plot* entre o evento de abertura do velo (s) e o primeiro segmento automático (em 4 segmentos) (s).

Modelo	R	R Square	Adjusted R Square
1	,981	,963	,962

Tabela 5.10: Sumário do Modelo obtido para o evento de abertura nasal e o primeiro segmento automático (em quatro segmentos)

A mesma sequência é apresentada para a relação existente entre o evento de fecho oral (ORALCLOSE) e o quarto segmento obtido de uma forma automática (SEG4.4) (figura 5.11, tabela 5.11).

Modelo	R	R Square	Adjusted R Square
1	,837	,700	,698

Tabela 5.11: Sumário do Modelo obtido para o evento de fecho oral e o quarto segmento automático.

A recta de regressão obtida para o conjunto de valores em análise é:

$$OralClose = 0,955 \times SEG4.4 - 0,029 \quad (5.4)$$

Os resultados apresentados nesta secção representam os valores de correlação e ajuste linear mais elevados que foram obtidos neste estudo.

Discussão

Os resultados apresentados anteriormente demonstram a existência de uma elevada correlação entre os eventos referidos, sendo estes, para a segmentação em três zonas estáveis,

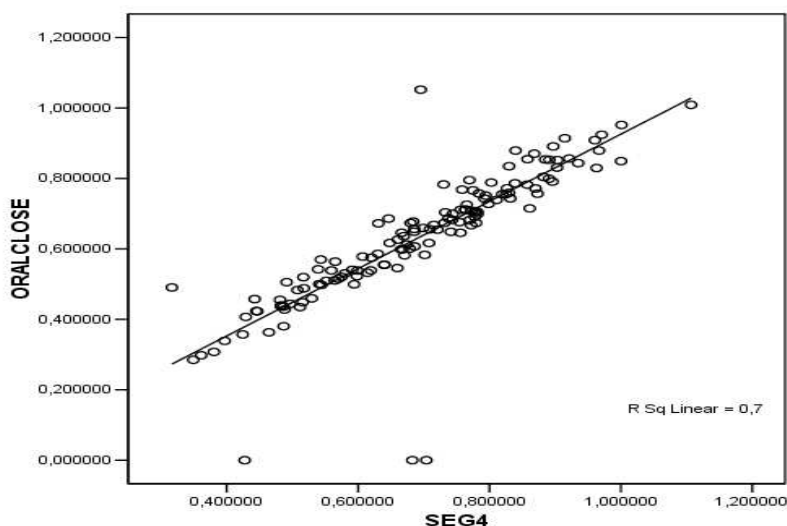


Figura 5.11: *Scatter Plot* entre o evento de fecho dos articuladores orais e o quarto segmento automático (em 4 segmentos).

NASALOPEN - SEG1 e ORALCLOSE - SEG2 e, para a segmentação em quatro zonas, os eventos NASALOPEN - SEG1.4 e ORALCLOSE - SEG4.4. Isto é válido, principalmente, para a representação do ajuste existente entre o primeiro segmento obtido por segmentação automática, em três e quatro zonas (SEG1 e SEG1.4) e o primeiro evento anotado no sensor do velo (NASALOPEN), que corresponde ao início da abertura deste e, consequentemente, ao fim da zona “oral” das vogais nasais.

Estes dados, embora preliminares, indicam a relevância de estudos mais específicos sobre este tipo de segmentação automática. Esta relevância torna-se mais evidente quando se compreende que este tipo de segmentação permitiria a obtenção de indirecta de eventos relacionados com o velo, essenciais para o estudo de vogais nasais, e consequentemente para a síntese articulatória.

5.3 Resumo

O capítulo apresenta a metodologia de análise e os resultados obtidos para a caracterização temporal das vogais nasais. O parâmetro *duração* foi analisado, quer para a totalidade da vogal nasal, quer para as suas diferentes fases.

No que diz respeito ao estudo da duração total foram apresentados valores médios, utilizando, quer os dados dos 25 informantes do *corpus* Regional, quer os dados provenientes do *corpus* recolhido através da tecnologia EMMA. Foram ainda listadas médias para cada vogal para, neste caso, as seis regiões contempladas no *corpus* Regional.

A informação proveniente dos sensores EMMA sobre os eventos de fecho e abertura do velo e dos articuladores orais, permitiu a obtenção de informação relevante sobre a duração

das várias fases das vogais nasais. Entre os vários valores obtidos destacam-se os apresentados para a duração do intervalo de abertura e de fecho do velo.

O estudo exploratório, cujo objectivo passa por obter informação sobre as diferentes fases da vogal, recorrendo apenas ao sinal acústico finaliza o capítulo.

Capítulo 6

Características Relacionadas com o Tracto

6.1 Introdução

Estudos anteriores de vogais nasais apontam no sentido de uma vogal nasal se realizar em três diferentes fases, sendo estas: uma vogal oral, seguida de uma breve fase transicional em que a nasalidade de sobrepõe à vogal, terminando num murmúrio nasal.

Pode supor-se que, sendo o véu palatino um articulador mais lento, a fase oral da vogal nasal se deveria a essa lentidão relativa, sendo necessário um tempo maior para que a abertura velo-faríngea alcance níveis suficientemente significativos para excitar os formantes e anti-ressonâncias do tracto nasal através da passagem de ar. Da mesma forma, a fase final do murmúrio pode dever-se a uma maior lentidão do véu palatino em voltar à posição de fecho da cavidade nasal.

Este capítulo pretende, assim, caracterizar a configuração adoptada pelo tracto na produção das vogais nasais e nas diferentes fases em que estas se realizam. Deste modo, começou por se efectuar o estudo relativo à fase inicial destes sons nasais (secção 6.2) e, deixando para trabalho posterior, a análise da fase medial, foi investigada a relação existente entre as características da fase final de uma vogal nasal e a consoante seguinte (secção 6.3).

6.2 Análise da Fase Inicial de uma Vogal Nasal no Contexto $C\tilde{V}C$

Associado a cada vogal existe um conjunto de formantes, ou ressonâncias, localizadas em frequências específicas. Os diferentes valores das formantes são determinados pelo tamanho e forma da câmara de ar na cavidade oral-faríngea. O movimento da língua (levantamento/abaixamento; avanço/retracção) pode resultar numa alteração da configuração desta câmara. Por exemplo, a elevação da língua está fortemente correlacionada com o decréscimo

de Primeira Formante (F1) e vice-versa. Também a retracção do corpo da língua está correlacionado com a diminuição da frequência de Segunda Formante (F2) (Sampson, 1999).

Nesta secção apresentam-se os resultados da análise da posição da língua na articulação das vogais orais e nasais em seis regiões de Portugal Continental, com base nos valores das duas primeiras formantes (Teixeira *et al.*, 2003a). Começa-se por apresentar informação acerca do processo de obtenção das formantes. É efectuado um estudo global, considerando os dados das várias regiões em conjunto, cujos resultados estão na forma de triângulos acústicos para cada sexo, para as vogais nasais e orais em separado. Também são apresentados gráficos de extremos e quartis (*boxplot*) para os valores de F1, F2 e F2-F1, para um estudo mais detalhado da altura e posição da articulação no eixo anterior-posterior.

Após esta análise em conjunto, repete-se a informação disponibilizada, de forma geral, para cada uma das regiões. Neste caso, acrescenta-se uma tabela com os triângulos acústicos dos quatro informantes de cada região.

A análise crítica detalhada dos resultados apresentados encerra esta secção.

6.2.1 Método

Extracção das formantes

O cálculo das formantes foi efectuado recorrendo ao *package* *Snack*¹ desenvolvido no *Kungliga Tekniska högskolan (Royal Institute of Technology)* (KTH), Suécia. A utilização deste pacote possibilita a obtenção de características do som que não seria possível obter recorrendo apenas ao Tcl/TK. O cálculo consiste essencialmente na execução dos dois seguintes comandos: **snacktrack::formant** *base.wav* **base.sfb** e **snacktrack::pitch** *base.wav* **base.sf0**

A estimativa das trajectórias das formantes é obtida recorrendo ao comando *formant*. Este comando possui um conjunto de parâmetros para os quais não é especificado qualquer valor sendo, portanto, utilizados os valores por defeito. Por exemplo, a opção **-numformants** que controla quantas formantes devem ser calculadas, e que no máximo toma o valor 7, está definida para tomar o valor 4 por defeito, sendo portanto esse o número de formantes calculadas neste caso. A opção **-framelength**, que especifica os intervalos entre os valores, toma o valor 0.001, enquanto que o parâmetro **-windowtype**, que é usado para seleccionar qual o tipo de janela aplicada, é por defeito o valor *Cos*⁴. O parâmetro **-windowlength** especifica o tamanho da janela em segundos e o seu valor por defeito é 0.049. Por seu lado, a opção **-preemphasisfactor**, que é usada para especificar qual a pré-ênfase aplicada ao sinal antes da aplicação da janela, toma o valor por defeito 0.7, enquanto que o parâmetro que especifica a ordem da análise LPC, **-lpcorder**, tem valor 12. A opção **-ds_freq**, que define a frequência de amostragem dos dados a serem usados na análise da frequência das formantes, é por defeito 1000 e, por último, o parâmetro **-nom_f1_freq** tem pré-definido o valor 500Hz por defeito. Este parâmetro especifica o valor nominal da frequência da primeira formante e o seu valor pré-definido assume que o tamanho do tracto vocal é de 17 cm e que a velocidade

¹<http://www.speech.kth.se/snack/>

do som é de 34000 cm/seg.

Os valores obtidos através deste método foram posteriormente tratados através do programa de análise estatística R e os respectivos resultados são apresentados de seguida.

Devido à extensão dos resultados obtidos, apenas serão apresentados nas secções seguintes os mais significativos, estando os restantes apresentados no Anexo F.

6.2.2 Resultados

Análise global dos dados

Como primeira análise, foram considerados os dados de todos os falantes das seis regiões contempladas. Para evitar o efeito género do informante, as análises foram efectuadas, sempre em separado, para os dados de informantes masculinos e femininos. Também se optou por separar, na grande parte das análises efectuadas, os resultados referentes às vogais orais dos resultados para as vogais nasais. No caso das nasais apenas foram consideradas para análise as que ocorrem depois de oclusiva (não nasal), por forma a que o seu início tenha as características de “vogal oral”, sendo aplicável o processo de extracção e análise das formantes aplicável às vogais orais. Como as análises efectuadas para as nasais só são possíveis para o seu início, a informação apresentada deve ser analisada tendo em conta este facto. Os resultados não são generalizáveis a toda a extensão da vogal.

Os gráficos seguintes (figura 6.1) apresentam a distribuição das duas primeiras formantes para a totalidade dos informantes. Os gráficos da esquerda representam a distribuição para os falantes masculinos e os superiores são relativos às vogais nasais.

Na figura 6.2 estão representados os diagramas de extremos e quartis para os valores de F1 em função da nasalidade e do género do informante. Uma representação equivalente é efectuada para os valores de F2, na figura 6.3 e para a diferença entre F2 e F1 são apresentados os valores na figura 6.4.

A análise efectuada para os valores de F1 e F2 em função da vogal é apresentada no Anexo F.

Análise por região

As Nasais

Nas figuras 6.5 e 6.6 apresentam-se os “triângulos acústicos” para as várias regiões contempladas no *corpus* recolhido. Na primeira figura, apresentam-se os resultados para as vogais nasais, na segunda, para as vogais orais. São apresentados resultados de forma mais detalhada, isto é, os resultados obtidos para cada uma das regiões de Portugal Continental, no Anexo F.

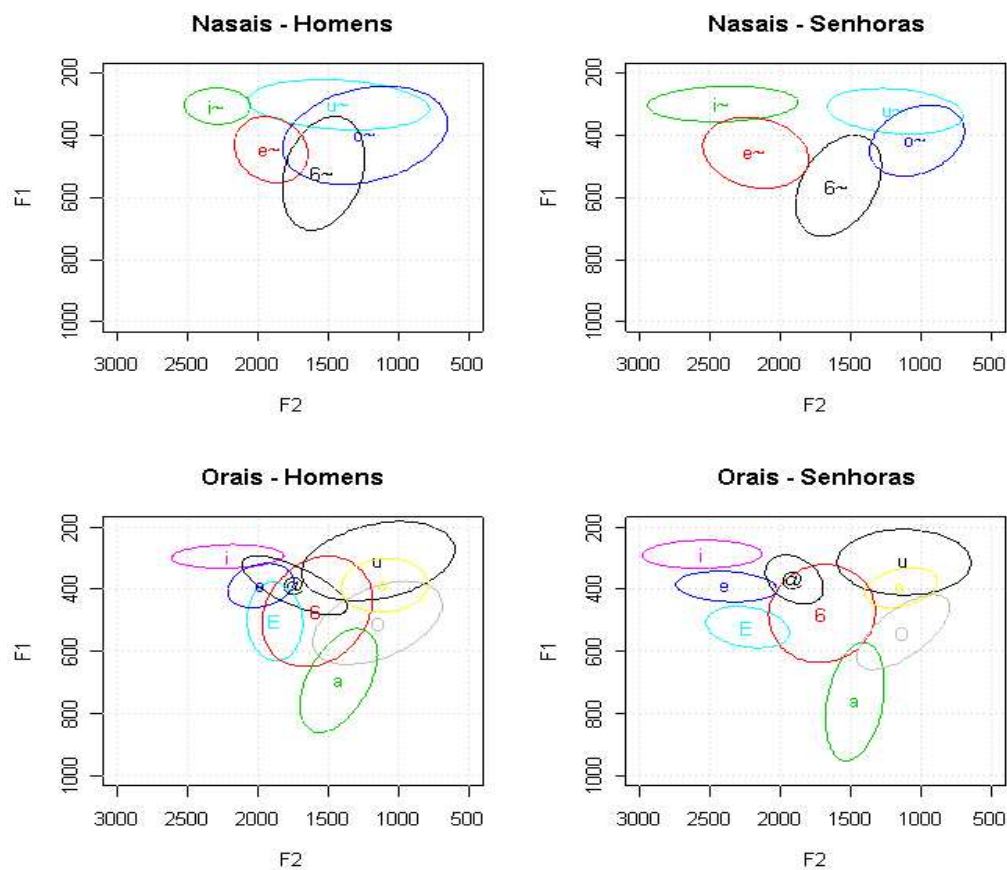


Figura 6.1: Representação da distribuição das duas primeiras formantes para a totalidade dos falantes, separando os valores pelo sexo (Homens do lado esquerdo) e existência ou não de nasalidade (nasais no topo) (Hz).

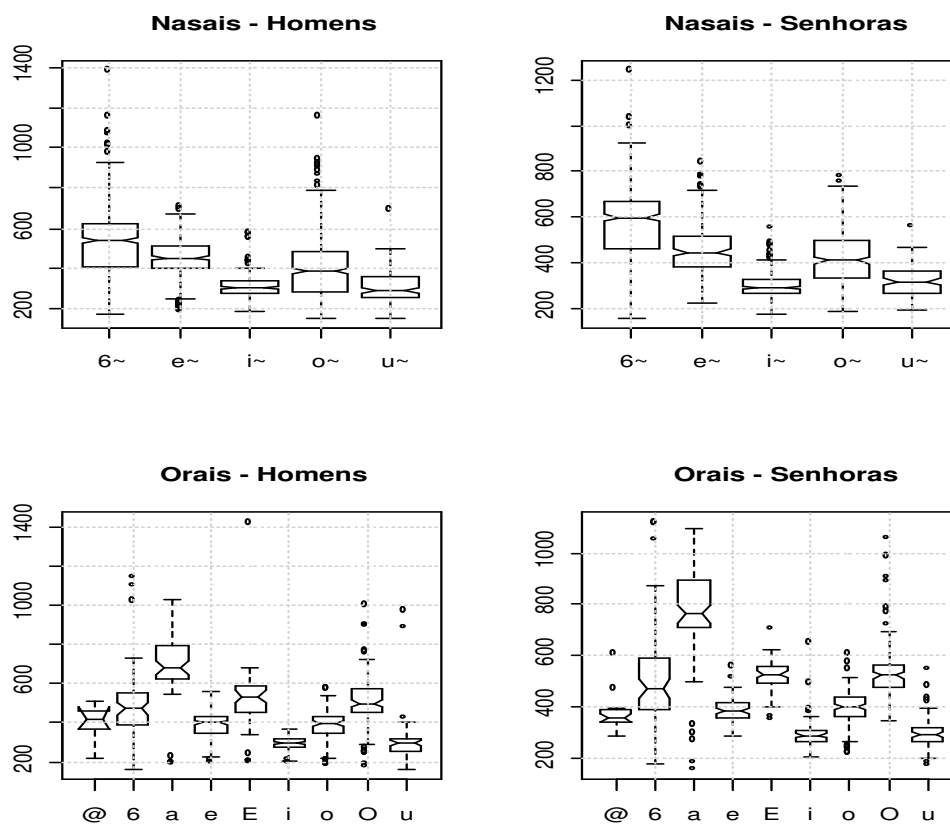


Figura 6.2: Boxplots de F1 (separação por sexo e nasalidade) (Hz).

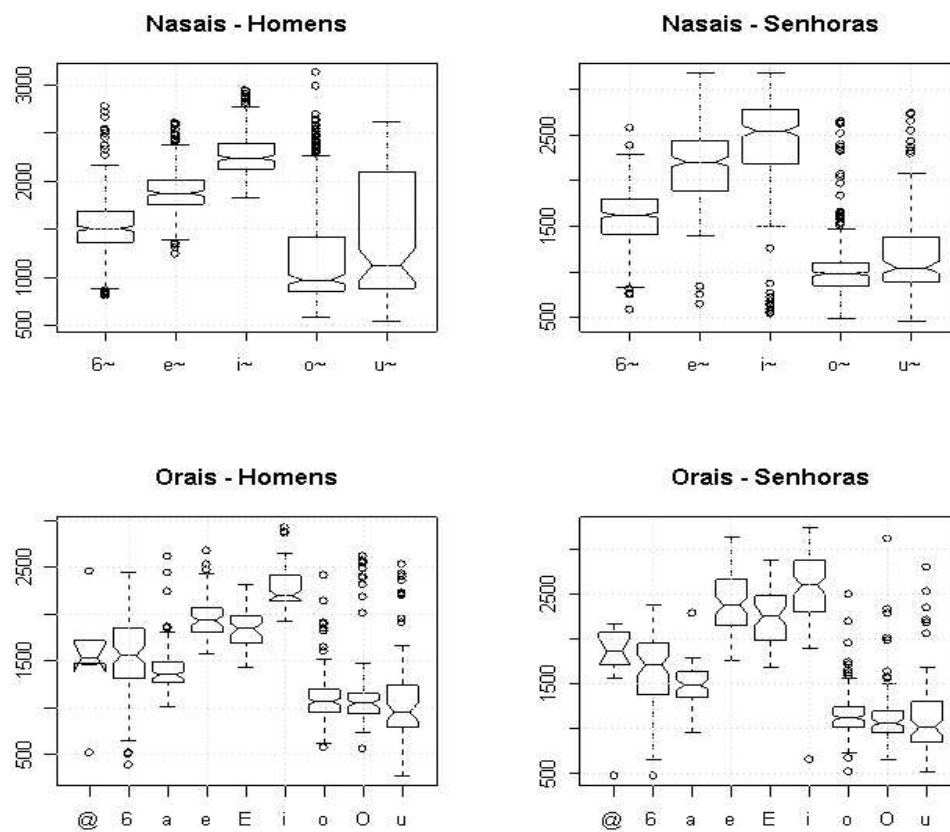


Figura 6.3: Boxplots de F2 (separação por sexo e nasalidade) (Hz).

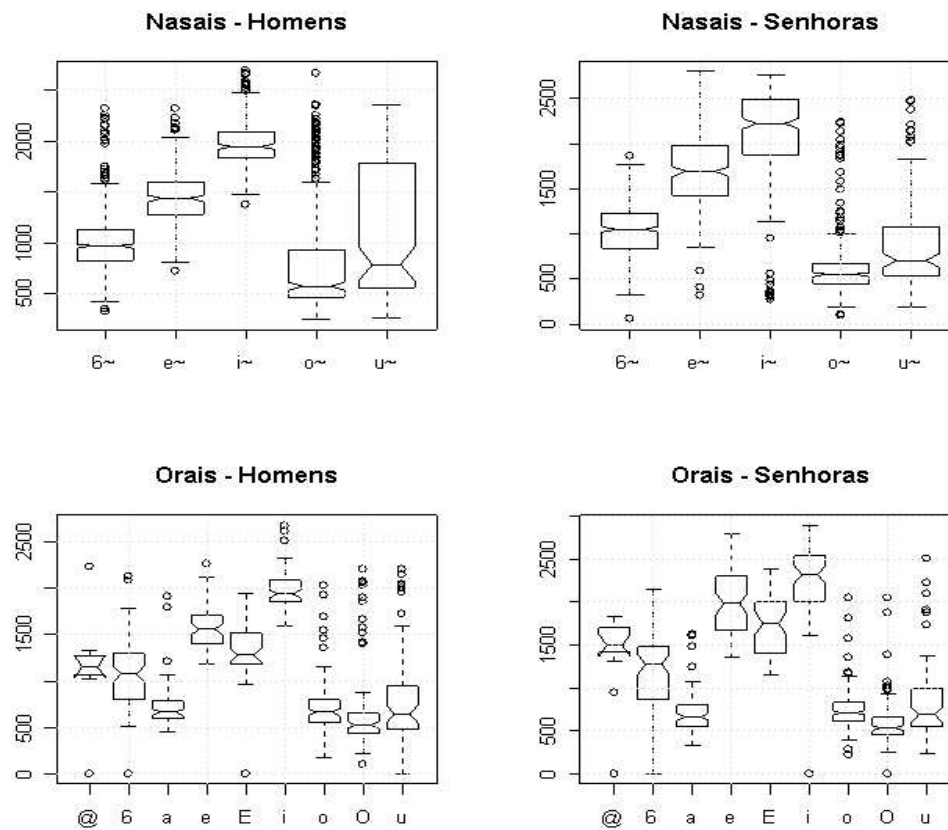


Figura 6.4: Boxplots da diferença entre F2 e F1 (separação por sexo e nasalidade) (Hz).

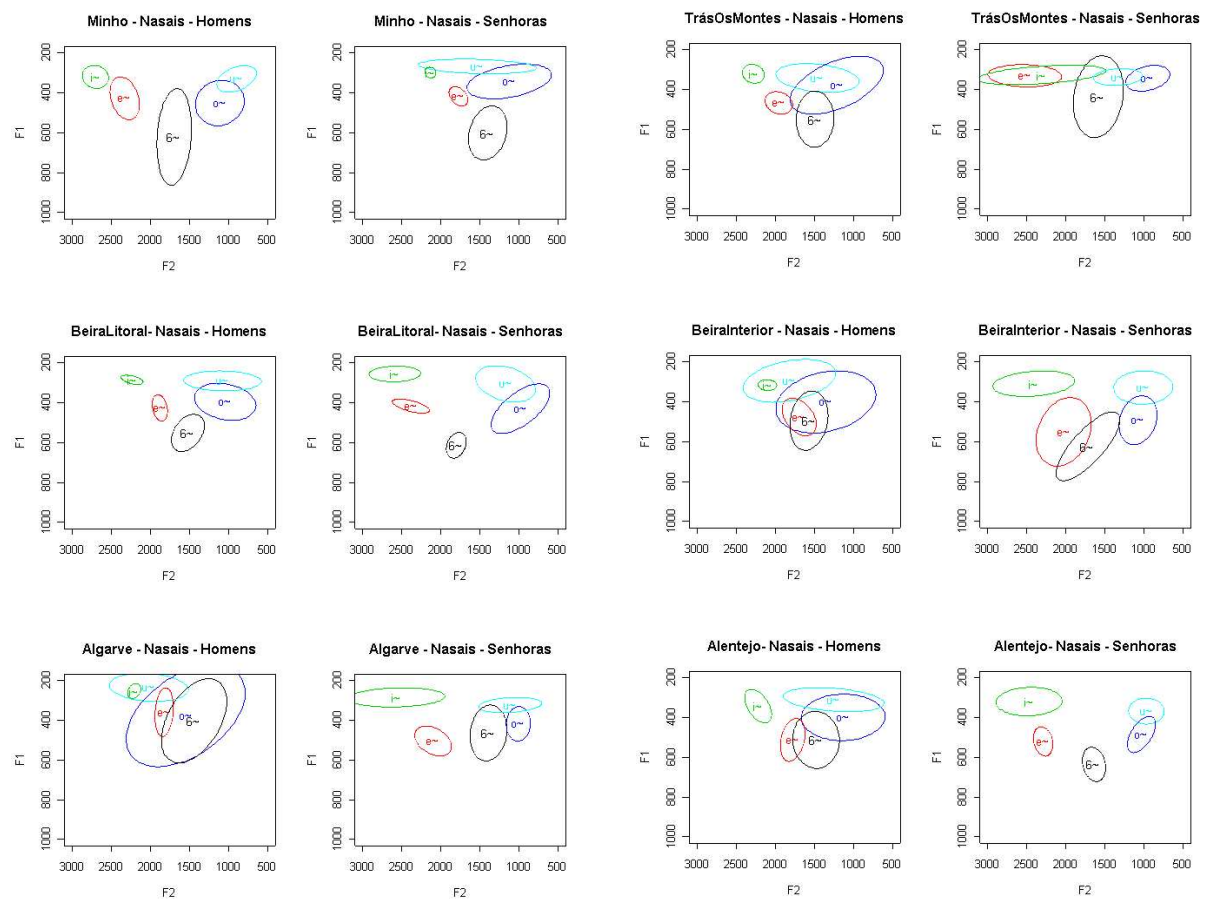


Figura 6.5: Triângulos acústicos para as vogais nasais das seis regiões contempladas no *corpus* (Hz).

As Orais

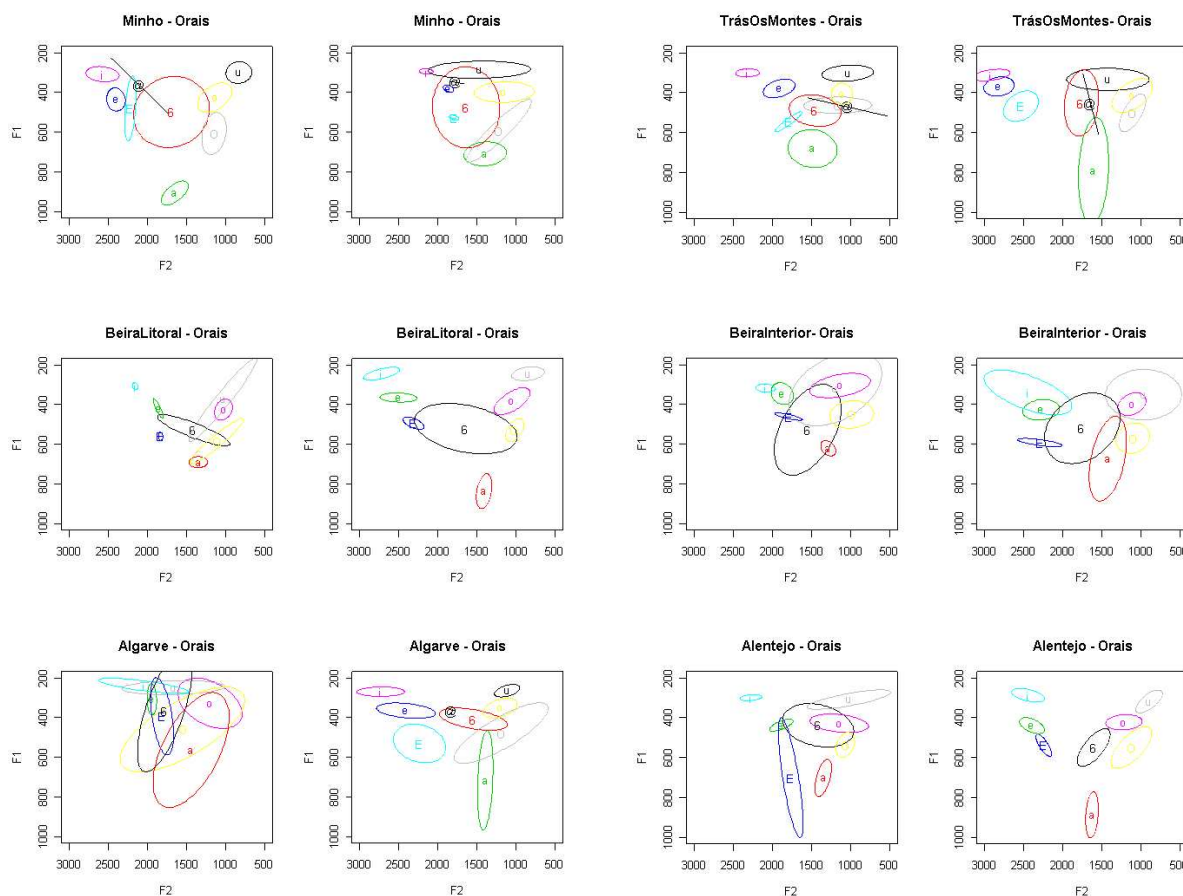


Figura 6.6: Triângulos acústicos para as vogais orais das seis regiões contempladas no *corpus* (Hz).

Região a Região

Os gráficos referentes à análise efectuada para cada uma das regiões estão apresentados no Anexo F. A informação é apresentada sob a forma de “triângulos acústicos” representando a distribuição das duas primeiras formantes para cada região.

6.2.3 Discussão

Todos os dados

Das figuras 6.1 e 6.2 (em especial desta última) ressalta que, quanto à altura da língua na produção das vogais para os falantes dos dois sexos:

1. a vogal nasal [6̃] tem uma altura mais próxima da vogal [6] do que do [a], situando-se, no entanto, entre estas duas vogais orais;

2. a vogal nasal [ẽ] pode considerar-se como próxima do [e];
3. a vogal nasal [õ] não apresenta grandes diferenças da vogal [o];
4. a vogal [ũ] é muito semelhante ao [u]
5. a vogal [ĩ] é muito semelhante ao [i].

Resumindo, o [ĩ] e [ũ] ocupam posições altas, o [ẽ] e [õ] posições médias (para alguns semi-fechadas) e o [ɐ̃] uma posição média um pouco mais baixa.

Das figuras 6.3 e 6.4 ressalta que, quanto à posição da língua, em termos do eixo anterior-posterior (zona de articulação) na produção das vogais para os falantes dos dois sexos:

1. a vogal nasal [ɐ̃] tem uma posição similar ao grupo [ə][ɐ][a], sendo mais aproximada das duas primeiras. Apesar de as diferenças não serem significativas, o valor médio da posição encontra-se entre os valores médios das vogais [a] e [ɐ];
2. a vogal [ẽ] tem uma posição similar ao grupo [e][E];
3. a vogal [õ] tem posição similar ao grupo [o][O][u];
4. a vogal [ũ] tem posição similar ao grupo [o][O][u];
5. a vogal [ĩ] tem posição similar ao [i], destacando-se pelo valor mais elevado de F2 em relação a todas as outras vogais.

Resumindo, quanto à zona de articulação surgem três grandes grupos, correspondentes à tradicional divisão em anteriores, centrais e posteriores. Não se torna necessária uma divisão diferente para as vogais nasais.

Da conjugação dos dois tipos de resultados anteriores, resulta que:

1. a vogal [ɐ̃], apesar de, numa primeira análise, mais próxima da oral [ɐ] apresenta uma configuração em termos de altura e zona de articulação intermédia entre as vogais orais [ɐ] e [a]. Não é tão baixa como [a] nem tão alta como [ɐ]. Também é mais central do que [a] sem ser tão central como o [ɐ] (veja-se Figura F.1 do Anexo F);
2. A vogal [ẽ] apresenta configuração comuns a ambas as orais [e] e [E], posicionando-se, em termos médios, entre estas duas vogais (veja-se Figura F.2 do Anexo F);
3. a vogal [ĩ] no seu início é muito aproximadamente produzida com a configuração oral da vogal oral [i] (veja-se Figura F.3 do Anexo F);
4. a vogal [õ] em termos de altura é mais próxima de [o], enquanto que em termos de zona de articulação se apresenta, em termos médios, ainda mais recuada que a vogal oral [O] (veja-se Figura F.4 do Anexo F);
5. a vogal [ũ] no seu início é muito aproximadamente produzida com a configuração oral da vogal oral [u] (veja-se Figura F.5 do Anexo F).

Região

Analisando as figuras 6.5 e 6.6 e todas as figuras seguintes (ver Anexo F), constata-se que, de uma maneira geral, o comportamento apresentado, para análise dos dados de todas as regiões em simultâneo, se repete com algumas pequenas diferenças pontuais, de que se destacam de seguida:

1. a grande sobreposição entre as configurações de [o] e [O], para os falantes de Trás-os-Montes, ser muito diferente do que ocorre, por exemplo, na Beira Litoral (Figuras F.10 e F.14);
2. a grande sobreposição entre [o] e [u] na Beira Interior (Figura F.18);
3. a produção do [ẽ] na Beira Interior e no Algarve com altura da língua correspondente à oral [E] (Figuras F.18 e F.26);
4. a grande sobreposição, mesmo em termos de altura, entre as configurações para as nasais [õ] e [ũ] por falantes do Alentejo (Figura F.22).

Resumindo, considera-se que existe variação em termos regionais das configurações utilizadas para a produção das vogais nasais e mesmo das orais.

6.3 Análise da Fase Final

6.3.1 Introdução

O processo articulatorio básico das vogais nasais, dos ditongos nasais e dos sons nasalizados tem, ao longo dos tempos, sido alvo de alguma controvérsia. No que diz respeito ao segmento final das vogais nasais surgiram várias interpretações sobre a relação entre este e a consoante seguinte. Algumas foram reunidas na revista do Laboratório de Fonética Experimental da Faculdade de Letras da Universidade de Coimbra, em 1966 (Lacerda e Head, 1966) e são apresentadas de seguida.

Oliveira Guimarães, no seu livro *Fonética Portuguesa, Compêndio da ortologia nacional* (citado em Lacerda e Head, 1966), refere que uma vogal nasal seguida de consoante oclusiva torna-se “impura”. Este aponta que tal vogal é seguida da ressonância nasal que acompanha a rapidíssima emissão de uma consoante nasal homorgânica da oclusiva seguinte, como por exemplo em: *lã tinta; lã cardada; lã branca* (pp. 43-44).

Em contradição surge a interpretação de J.I. Louro em *Estudo e Classificação das Vogais*, p.242 (citado em Lacerda e Head, 1966):

“As chamadas *vogais nasais* (como os *ditongos nasais*) são propriamente *oro-nasais*, visto que a ressonância nasal (devida como se sabe à comunicação do canal sonoro com as fossas nasais por abaixamento do véu palatino) se junta à

vogal emitida normalmente pela boca. Só a vogal @~ [a~] pode ser *oro-nasal* ou exclusivamente *nasal*, neste caso emitida apenas pelas fossas nasais (com oclusão da boca ou da região velar). As outras vogais ([b~], [e~], [i~], [o~], [u~]), quando mediais ou no interior das frases, são geralmente ligadas (ou mesmo substituídas na sua parte final) por um [a~] (formando com elas uma espécie de discreto ditongo crescente). São as vibrações deste [a~] que, nos gráficos, fazem pensar na existência, em português, de verdadeiras consoantes nasais, em fim de sílaba interna”.

No estudo das vogais nasais, convém distinguir as seguintes posições e contextos:

1. final;
2. antes de consoante oclusiva;
3. antes de consoante restritiva.

Relativamente a análise da posição final é de considerar os resultados a que chegaram vários investigadores, entre os quais figura Morais-Barbosa com o artigo *Les voyelles nasales portugaises: interprétation phonologique*, p.692 (citado em Lacerda e Head, 1966). Este refere que, na presença de uma vogal nasal em posição final, e após a emissão da vogal, existe uma oclusão da passagem oral, que tem lugar na região velar. Morais-Barbosa descreve a formação de uma consoante nasal, que denomina por [J] e indica a presença desta em palavras tais como [lɐ~J] e [fi~J]. Este refere ainda que esta consoante não é perceptível por ouvintes não treinados foneticamente, mas refere a sua presença nos traçados obtidos através do quimógrafo e espectógrafo.

A mesma interpretação foi publicada por R. de Sá Nogueira, em *Elementos para um Tratado de Fonética Portuguesa*, p.24 (obra citada em Lacerda e Head, 1966). Este refere que quando uma vogal nasal é proferida com o véu palatino primeiramente na posição de indiferença e depois adaptado ao pós-dorso da língua, em vez de uma vogal puramente nasal, o que se obtém são dois fonemas distintos: uma vogal nasal e o n velar. Este exemplifica o exposto utilizando a palavra *irmã* e as suas distintas pronúncias *irmã*, *irmãn* e *irmâ* com o *a* desnasalado.

No que refere à posição da vogal nasal antes de oclusiva, Morais-Barbosa (obra referida anteriormente, pp.692-693) afirma que entre uma vogal nasal e uma consoante oclusiva existe uma consoante nasal nítida. Este refere ainda que o ponto de articulação de tal consoante depende do ponto de articulação da oclusiva. Assim, caso a vogal se encontre antes de oclusiva labial (p, b) ocorre uma consoante labial nasal [m], no caso de após a vogal nasal existir uma oclusiva dental (t, d) a consoante nasal intermediária é uma dental nasal [n], e por último, este afirma que, antes de uma oclusiva velar existe um [J] velar. O autor afirma mesmo que nunca encontrou qualquer exemplo de vogal nasal seguida de oclusiva oral sem a presença de uma oclusiva nasal intermediária.

Esta interpretação está presente nas regras de transcrição propostas por Lacerda e G. Hammarström (*Transcrição Fonética do Português Normal*, p.123) (citado em Lacerda e Head, 1966). Os seguintes exemplos, citados por Lacerda e Head (1966), ilustram tais regras:

Gonçalves Viana em *Exposição*, pp.52-53, observou o seguinte:

“Quando a uma vogal nasal se segue consoante explosiva, além dessa vogal nasal, ouve-se atenuada, reduzida, uma consoante nasal, homorgânica com essa explosiva
....”

A citação de Oliveira Guimarães, apresentada no início desta secção, demonstra a concordância deste autor com as palavras de Gonçalves Viana.

No entanto, J.I.Louro surge com uma interpretação oposta (*ob. cit.*, p.242):

“...São as vibrações deste [a~] que, nos gráficos, fazem pensar na existência, em português, de verdadeiras consoantes nasais, em fim de sílaba interna. No entanto, estas consoantes nasais à semelhança do que se passa em francês, devem considerar-se meramente gráficas, quer sejam seguidas de consoante oclusiva, quer de constrictiva. É que, embora escrevendo-se na actual ortografia, em nenhum caso há consciência de se pronunciarem e, na realidade, também não lhes correspondem quaisquer movimentos articulatorios activos, próprios (compare-se *mata* e *manta* ou *lida* e *linda* ou notem-se as diferenças, por exemplo, entre *manta*, *malta* e *marta* ou *ganço* e *garça*). As vibrações dos gráficos, como dissemos correspondem apenas a um [a~] de ligação (ou mesmo de substituição parcial). É certo que, especialmente no caso da *nasal* ser seguida de consoante labial, pode surgir o equivalente fonético e articulatorio da consoante nasal, do *m*. Mas este *m* é também inconsciente e passivo, resultando apenas do mesmo [a~] e da oclusão da boca para a pronúncia da labial seguinte (compare-se *lã* e *lamber* ou *lã preta*...).”

Por último, e no que diz respeito ao caso de vogal nasal seguida de consoante restritiva, surgem ainda mais discordâncias. Gonçalves Viana (*Portugais*, p.16), bem como H. Sten (*Les Particularités de la Langue Portugaise*, p.32) referem a existência de vogais nasais puras apenas quando em posição final (contrariando o que se disse anteriormente relativamente a vogais nasais nessa posição) ou quando são seguidas de *consoantes contínuas*.

O estudo de Morais-Barbosa (*ob. cit.*, p.703) relata outra interpretação. Este entende que existe sempre uma consoante nasal, mais ou menos curta ([m], [n], [J]), antes de uma oclusiva e que por vezes também existe antes de uma consoante restritiva ou uma pausa.

Interessa ainda considerar a opinião de Sá Nogueira (*ob. cit.*, pp. 71-72) que indica a presença de consoantes nasais entre uma vogal e a oclusiva seguinte. Este diz que, na sequência do tipo vogal mais consoante nasal seguida de consoante (*ob. cit.*, p.117), “As consoantes nasais que podem no português normal encontrar-se nesta situação são [m],[n],[J]. Neste caso a nasalação da vogal precedente é nítida”.

É de notar que esta interpretação considera a vogal nasal como um exemplo de “assimilação”, tratando-se, deste modo, de uma vogal *nasalizada*.

Estudos recentes (como o de de Sousa (1994)) indicam que as vogais nasais se realizam em três momentos relativamente distintos, a saber: uma vogal oral, seguida depois por uma breve fase transicional em que a nasalidade se sobreporia à vogal, terminando por fim num murmúrio nasal. Esta conclui que o murmúrio é bastante coarticulado à vogal, não apresentando sinais aparentes de transições para um ponto de articulação consonantal.

Caso estas observações sejam confirmadas, estaria assim descartada a hipótese da realização de uma consoante nasal [J] em posição anterior a uma oclusiva velar (como na palavra “canga”, por exemplo). O que existiria, foneticamente, seria um murmúrio coarticulado à vogal nasal, e sem existência consonantal autónoma em relação a esta vogal.

É para o esclarecimento destas duas teorias que se pretende contribuir nas secções seguintes.

6.3.2 Objectivos

Como se pode verificar do exposto anteriormente, não é claro nos resultados apresentados por outros investigadores até que ponto as características finais das vogais nasais são essencialmente dependentes da co-articulação com a consoante seguinte ou se dependem apenas das suas características intrínsecas. Assim, e de modo a tentar esclarecer o problema apresentado, foi desenvolvido um processo cujo objectivo passa por obter as distâncias espectrais entre os segmentos nasais em análise.

6.3.3 Método

Com o intuito referido, foram obtidas as características espectrais no final das vogais nasais e compararam-se com as características espectrais das consoantes nasais e de outras vogais nasais. A semelhança espectral foi aferida através de medidas que tentam traduzir uma distância espectral.

De seguida são apresentadas algumas destas medidas.

Medidas de Distância

Em processamento de voz, defende-se que uma medida de distância entre dois segmentos de sinais de voz x e y

$$d : (x, y) \in E \times E \rightarrow d(x, y) \in \mathbb{R}_0^+$$

deve satisfazer, pelo menos, as seguintes propriedades (Deller *et al.*, 1993); (Gray Jr. e Markel, 1976):

1. $d(x, y) = d(y, x)$ (simetria)

2. $d(x, y) > 0 \forall x \neq y$
 $d(x, x) = 0$ (definida positiva)
3. $d(x, y)$ deve ter uma interpretação com significado físico no domínio da frequência
4. Deve ser possível avaliar eficientemente $d(x, y)$

Distância espectral

Uma medida de distância espectral é então a representação numérica, não negativa, da semelhança entre espectros de dois sinais. Deste modo, e de acordo com as propriedades anteriores, dois sinais iguais têm uma distância espectral nula e esta medida aumenta à medida que os espectros de tais sinais diferem.

Existem várias medidas de distância espectral amplamente usadas em processamento de voz. Neste estudo usaram-se medidas baseadas em distâncias euclidianas aplicadas, quer a vectores de coeficientes cepstrais em escala Mel (*Mel Frequency Cepstral Coefficients* (MFCC)), quer a vectores de coeficientes *Perceptual Linear Prediction* (PLP) (Hermansky e Morgan, 1994); (Hermansky, 1990). Para além destas, foram calculadas medidas com base no espectro de potência, como é o caso das distâncias de Itakura, Itakura-Saito e Itakura-Saito simétrica (Rabiner e Juang, 1993).

As medidas referidas são apresentadas nas secções seguintes.

Medidas baseadas em distâncias euclidianas

Uma das medidas mais utilizadas em processamento de voz, é a distância euclidiana de vectores de coeficientes cepstrais. Sejam x e y dois vectores contendo os coeficientes calculados nos instantes de interesse (neste caso, na parte final de uma vogal nasal e na parte medial de uma consoante nasal). A distância euclidiana no espaço \mathbb{R}_k dos coeficientes cepstrais, ou distância cepstral truncada (*Truncated Cepstral Distance*), como também é designada, é dada por:

$$d_c^2 = \sqrt{\sum_{k=1}^p (c_k - c'_k)^2} \quad (6.1)$$

As distâncias euclidianas calculadas neste trabalho, isto é, as distâncias entre vectores de MFCC e de coeficientes PLP, serão designadas neste texto por $dMFCC$ e $dPLP$ e definidas pelas equações:

$$dMFCC = \sqrt{\sum_{k=1}^p (m_k - m'_k)^2} \quad (6.2)$$

$$dPLP = \sqrt{\sum_{k=1}^p (l_k - l'_k)^2} \quad (6.3)$$

onde p é o número de coeficientes e m_k, m'_k e l_k, l'_k são, respectivamente os coeficientes cepstrais de Mel e PLP.

Medidas com base no espectro de potência

Considerando os espectros de potência dos sinais a comparar, é possível obter outro tipo de distâncias. Uma destas medidas de distância espectral foi desenvolvida por Itakura, em 1975, e é conhecida como razão de verosimilhança logarítmica (*log likelihood ratio*). Define-se como sendo:

$$dIT(S_1, S_2) = \log \left(\int_{-\pi}^{\pi} \frac{S_1(w)}{S_2(w)} \frac{dw}{2\pi} \right) - \int_{-\pi}^{\pi} \log \left(\frac{S_1(w)}{S_2(w)} \right) \frac{dw}{2\pi} \quad (6.4)$$

onde S_1 e S_2 são os espectros de potência dos sinais em análise.

Sendo $V(w) = \log S_1(w) - \log S_2(w)$, a distância espectral de Itakura-Saito é definida por:

$$dIS(S_1, S_2) = \int_{-\pi}^{\pi} \left(e^{V(w)} - V(w) - 1 \right) \frac{dw}{2\pi} \quad (6.5)$$

Observa-se em (6.5) que $dIS(S_1, S_2) \neq dIS(S_2, S_1)$, isto é, a distância de Itakura-Saito é uma medida não simétrica. Para contornar esta situação é, por vezes, utilizada uma outra medida designada Itakura-Saito simétrica ou de coseno hiperbólico, que é definida em (6.6).

$$\begin{aligned} dCOSH(S_1, S_2) &= \frac{dIS(S_1, S_2) + dIS(S_2, S_1)}{2} = \\ &= \int_{-\pi}^{\pi} \left(e^{V(w)} - V(w) - 1 + e^{-V(w)} + V(w) - 1 \right) \frac{dw}{2\pi} = \\ &= \int_{-\pi}^{\pi} \left(\cosh V(w) - 1 \right) \frac{dw}{2\pi} \end{aligned} \quad (6.6)$$

Método de cálculo das distâncias espectrais

A implementação destas medidas foi baseada em vários pacotes desenvolvidos para o MATLAB (Little e Shure, 1992).

Para obter os vectores de MFCCs foi utilizada a *Auditory Toolbox*² desenvolvida na *Interval Research Corporation*. Estes coeficientes foram obtidos recorrendo à rotina *mfcc* do referido pacote, que aplica janelas de Hamming de 20ms ao sinal de entrada (vogal ou consoante nasal) e utiliza um banco de filtros construído através 13 filtros linearmente espaçados, seguido de 27 filtros com um espaçamento logarítmico.

Foram desta forma obtidos 13 coeficientes cepstrais, sendo o primeiro, C_0 , função da energia do sinal e os restantes coeficientes, $C_1 - C_{12}$, calculados de modo a que o seu valor esperado seja nulo.

Foi referida anteriormente a utilização de uma abordagem baseada em *Perceptual Linear*

²<http://svr-www-eng.cam.ac.uk>

Prediction. Esta abordagem foi implementada recorrendo à rotina *rastaplp* disponível na ferramenta *RastaMat*³. Esta foi utilizada para a obtenção de 12 coeficientes PLP, através da conversão da onda sonora numa sequência de *feature frames*. As medidas dMFCC e dPLP foram calculadas recorrendo ao pacote de processamento de voz para MatLab, *VoiceBox*⁴.

Cada conjunto de coeficientes obtido foi organizado em 7 matrizes correspondentes a cada um dos fonemas nasais analisados, isto é, cinco vogais ([6̃], [ẽ], [ĩ], [õ] e [ũ]) e duas consoantes nasais ([m] e [n]). Posteriormente, a rotina *dsteusq*, que calcula a distância euclidiana entre todos os pares de linhas de duas matrizes, foi utilizada para se obter as distâncias entre vogais iguais, entre diferentes vogais nasais e também entre as consoantes nasais e cada uma das vogais nasais consideradas.

Para o cálculo das dIT, dIS e dCOSH foi utilizada uma implementação baseada no pacote Voicebox referido anteriormente. As rotinas *distitar*, *distisar* e *distchar* calculam as distâncias espectrais de Itakura, Itakura-Saito e COSH entre conjuntos de coeficientes LPC. Deste modo, realizou-se uma análise LPC de ordem 12, utilizando para o efeito, janelas de Hamming com 20ms de comprimento.

Os coeficientes necessários para obter cada uma das medidas referidas foram calculados sobre segmentos de sinal de voz correspondentes à zona de interesse. Uma vez que se pretende avaliar qual a relação existente entre o segmento final das vogais nasais e a consoante que lhes segue, foram obtidos os coeficientes necessários para o cálculo das distâncias espectrais, para cada uma das cinco vogais nasais, a cerca de 90% da sua duração total e para a zona medial das consoantes nasais.

Deste modo, e de modo a poder efectuar-se algumas comparações, foram calculadas distâncias para os segmentos correspondentes a vogais iguais (por exemplo, a distância entre os espectros dos segmentos finais, tal como a distância entre os espectros das nasais [6̃]), distâncias espectrais de vogais diferentes (por exemplo, distância entre os espectros de vogais [6̃] e todas as outras ([ẽ], [ĩ], [õ] e [ũ])) e, por último, distâncias espectrais entre o segmento final de cada uma das vogais e consoantes nasais (exemplo, distância entre os espectros do segmento final das nasais [6̃] e das consoantes [m] e [n]).

Cada uma destas distâncias foi denominada, respectivamente, de:

- $V_i V_i$ - distância entre os espectros de vogais nasais iguais;
- $V_i V_j$ - distância entre os espectros de vogais diferentes;
- $V_i C_k$ - distância entre os espectros de vogais nasais e consoantes nasais,

onde $i=1(=[6̃]), 2(=[ẽ]), 3(=[ĩ]), 4(=[õ]), 5(=[ũ])$, $j=1(=[6̃]), 2(=[ẽ]), 3(=[ĩ]), 4(=[õ]), 5(=[ũ])$ e $j \neq i$, e $k=1(=[m]), 2(=[n])$.

Os resultados obtidos são apresentados na secção seguinte.

³<http://labrosa.ee.columbia.edu/matlab/rastamat>

⁴<http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox>

6.3.4 Resultados

Os valores obtidos resultam do processamento estatístico (efectuado com o SPSS) de um elevado conjunto de valores correspondentes a distâncias espectrais. Estas foram processadas de acordo com o contexto em que se encontravam, sendo estes antes de consoante oclusiva surda ou sonora (denominados no restante documento de ptk e bdg, respectivamente).

Este processamento estatístico resultou num enorme conjunto de valores correspondentes a valores originados por testes estatísticos, tais como testes ANOVA, testes de homogeneidade e testes post-hoc de Tukey (Guimarães e Cabral, 1997).

De modo a tornar os dados obtidos legíveis optou-se por sistematizá-los em tabelas (tabelas 6.1, 6.2, 6.3, 6.4 e 6.5) que indicam de uma forma qualitativa quais os valores atribuídos por cada tipo de distância espectral utilizada. A codificação utilizada é a seguinte:

- valor **1** indica que foi atribuído ao grupo (V_iV_i , V_iV_j e V_iC_k) em análise o **menor** valor de distância espectral;
- valor **2** indica que foi atribuído ao grupo em análise o valor de distância espectral **intermédio**;
- valor **3** indica que foi atribuído ao grupo em análise o **maior** valor de distância espectral.

Contexto	V_iC_k	V_iV_i	V_iV_j	Tipo
	2	1	3	dMFCC
	2	3	1	dCOSH
	3	1	2	dITA
	3	2	1	dITAS
	3	1	2	dPLP
bdg	2	1	3	dMFCC
	2	3	1	dCOSH
	3	2	1	dITA
	3	2	1	dITAS
	3	2	1	dPLP
ptk	2	1	3	dMFCC
	3	2	1	dCOSH
	3	1	2	dITA
	3	2	1	dITAS
	3	1	2	dPLP

Tabela 6.1: Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [6~].

Na tabela 6.1, que apresenta os resultados obtidos para a vogal [6~], evidencia-se o facto de apenas as medidas dMFCC, dITA e dPLP associarem a distância entre os espectros de vogais iguais (neste caso, V_iV_i) à menor distância. Seria, obviamente, de esperar que os valores retornados por todas as medidas indicassem esta distância como menor e por isso só serão

consideradas para discussão os resultados provenientes de medidas de distância espectral que consigam fazer esta distinção.

Deste modo, das três medidas em consideração (dMFCC, dITA e dPLP) apenas a distância calculada a partir dos coeficientes Mel indica como menor a distância entre os espectros dos segmentos finais das vogais e as consoantes nasais.

Os outros resultados parecem indicar que o segmento final das nasais está mais próximo em termos espectrais do segmento final de outras nasais, isto é, que as características finais das vogais nasais são intrínsecas.

Contexto	V_2C_k	V_2V_2	V_2V_j	Tipo
	1	2	3	dMFCC
	2	1	3	dCOSH
	3	2	1	dITA
	1	2	3	dITAS
	3	1	2	dPLP
bdg	2	1	3	dMFCC
	2	1	3	dCOSH
	3	2	1	dITA
	2	1	3	dITAS
	3	1	2	dPLP
ptk	2	1	3	dMFCC
	2	1	3	dCOSH
	3	1	2	dITA
	1	2	3	dITAS
	3	1	2	dPLP

Tabela 6.2: Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [ẽ].

Da análise dos resultados obtidos para a nasal [ẽ], listados na tabela 6.2, é possível verificar que apenas as medidas dCOSH e dPLP indicam a distância V_2V_2 como menor. No que diz respeito aos valores atribuídos por estas medidas aos restantes segmentos em análise verifica-se uma clara contradição. A dCOSH relaciona o segmento final das nasais, primeiramente, com as consoantes nasais, enquanto que a medida dPLP o faz, em primeiro lugar, com o segmento final de outras vogais.

O resultado obtido com as medidas dPLP e dCOSH para a nasal [ẽ] é também verificado para a vogal [ĩ] (tabela 6.3), bem como para as nasais [õ] (tabela 6.4) e [ũ] (tabela 6.5).

A tabela 6.4 tem a particularidade de indicar, para todas as medidas de distância espectral utilizadas, o menor valor para as distâncias entre vogais iguais. No entanto, o facto de tal não ter acontecido com as restantes vogais nasais deve ser considerado.

Contexto	V_3C_k	V_3V_3	V_3V_j	Tipo
	1	2	3	dMFCC
	2	1	3	dCOSH
	3	2	1	dITA
	2	1	3	dITAS
	3	1	2	dPLP
bdg	1	2	3	dMFCC
	2	1	3	dCOSH
	3	1	2	dITA
	2	1	3	dITAS
	3	1	2	dPLP
ptk	2	1	3	dMFCC
	3	2	1	dCOSH
	3	2	1	dITA
	3	2	1	dITAS
	3	1	2	dPLP

Tabela 6.3: Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [ĩ].

Contexto	V_4C_k	V_4V_4	V_4V_j	Tipo
	2	1	3	dMFCC
	2	1	3	dCOSH
	3	1	2	dITA
	2	1	3	dITAS
	3	1	2	dPLP
bdg	2	1	3	dMFCC
	2	1	3	dCOSH
	3	1	2	dITA
	2	1	3	dITAS
	3	1	2	dPLP
ptk	2	1	3	dMFCC
	3	2	1	dCOSH
	3	2	1	dITA
	3	2	1	dITAS
	3	1	2	dPLP

Tabela 6.4: Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [õ].

Contexto	V_5C_k	V_5V_5	V_5V_j	Tipo
	1	2	3	dMFCC
	2	1	3	dCOSH
	3	2	1	dITA
	2	1	3	dITAS
	3	1	2	dPLP
bdg	1	2	3	dMFCC
	2	1	3	dCOSH
	3	1	2	dITA
	2	1	3	dITAS
	3	1	2	dPLP
ptk	2	1	3	dMFCC
	3	2	1	dCOSH
	3	1	2	dITA
	3	2	1	dITAS
	3	1	2	dPLP

Tabela 6.5: Resumo qualitativo dos valores atribuídos por cada medida de distância espectral utilizada, para a vogal nasal [ũ].

6.3.5 Discussão

Os resultados obtidos apresentam claras contradições entre si. A tabela 6.6 apresenta o número de vezes que cada medida utilizada atribuiu à distância V_iV_i (distância entre vogais iguais) o maior, o menor e o valor intermédio (representados respectivamente por 3, 1 e 2).

	Distâncias		
V_iV_i	1	2	3
dMFCC	2	3	0
dCOSH	4	0	1
dITA	2	3	0
dITAS	3	2	0
dPLP	5	0	0

Tabela 6.6: Resumo quantitativo dos valores atribuídos por cada medida de distância espectral à distância entre a mesma vogal (V_iV_i).

Da tabela 6.6 é claro que apenas a distância dPLP considerou esta distância como inferior em todas as situações. A dCOSH também se destaca, tendo atribuído em 90% das vezes o menor valor a esta distância. Uma vez que se espera que esta distância seja, realmente, a menor, a discussão seguinte terá em consideração estes dados não considerando as restantes medidas.

Assim, existe uma tendência clara para que a dPLP classifique como menor a distância existente entre as mesmas vogais (V_iV_i), logo seguida da distância entre uma vogal e outras vogais nasais (V_iV_j). Parece, assim, indicar que os espectros do segmento final das vogais nasais se assemelham mais entre si, do que aos espectros das consoantes nasais, corroborando

desta forma a teoria que defende que as características finais das vogais nasais são intrínsecas.

A dCOSH por seu lado inverte os resultados anteriores, classificando a distância existente entre diferentes vogais nasais como superior à distância entre vogais nasais e consoantes nasais.

Tal contradição impede a obtenção de qualquer conclusão. No entanto, a utilização de medidas de distâncias espectral baseadas em percepção (referidas por De (1993) e Mannell (1994)) (fora do limite temporal deste projecto) poderá contribuir para o esclarecimento deste assunto.

6.4 Resumo

Na análise das características do tracto durante a produção das várias vogais nasais, convém distinguir três diferentes fases, sendo estas:

- fase inicial;
- fase medial;
- fase final.

É o estudo de duas destas fases que preenche este capítulo.

A análise da fase inicial das vogais nasais foi efectuada usando a relação entre F1, F2 e a altura da vogal e zona de articulação, recorrendo à informação proveniente do *corpus* Regional. O facto de este *corpus* contemplar dados provenientes de várias regiões do país, permitiu obter informação relevante e pouco explorada.

Deixando para mais tarde o estudo da parte medial, foi realizada a análise da fase final das vogais nasais, que se baseou na aplicação de medidas de distâncias espectrais para aferir a semelhança entre o segmento final destas e a consoante seguinte. Com este estudo pretendeu-se contribuir para uma controvérsia existente entre os investigadores da área, onde de um lado estão os que defendem que entre uma vogal nasal e a consoante oclusiva seguinte existe, nitidamente, uma consoante nasal, e do outro estão os que referem que o murmúrio nasal existente na fase final das vogais nasais é consideravelmente coarticulado à vogal e não apresenta sinais aparentes de transições para um ponto de articulação consonantal. Os resultados, que se devem considerar preliminares, não foram no entanto conclusivos.

Capítulo 7

Conclusões

7.1 Resumo do Trabalho

A dissertação apresentada neste documento tem como objectivo principal contribuir para a caracterização das vogais nasais existentes no PE. O trabalho surge de uma colaboração efectuada com o Departamento de Línguas e Culturas - CLC da Universidade de Aveiro, no âmbito do projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português”, cuja ideia passa por adquirir conhecimento mais aprofundado sobre estes sons nasais para a produção das diferentes vogais nasais, utilizando, para tal, o sintetizador articulatório em desenvolvimento nesta Universidade.

Deste modo, os estudos efectuados neste trabalho abrangem diversas áreas, como, por exemplo, a caracterização da fonte glotal durante a produção dos sons nasais e a caracterização temporal destes, contemplando aspectos pouco analisados anteriormente, como a variação regional. Assim, foram analisados parâmetros como a F0, o OQ, *jitter* e a duração (total e das diferentes fases) para cada vogal nasal. Efectuaram-se comparações com os valores obtidos para as vogais orais, e entre falantes, regiões e outros. Um estudo com o objectivo de obter informação sobre as várias fases de uma vogal nasal, tendo por base apenas o sinal acústico, foi efectuado. O interesse deste estudo tem origem na dificuldade de aplicação da tecnologia EMMA a um conjunto alargado e aleatório de pessoas e baseia-se na exploração de correlações existentes entre as várias fases que constituem uma vogal nasal e as zonas com características espectrais estáveis.

As características das vogais nasais relacionadas com o tracto foram também analisadas, tendo este estudo incidido sobre o início e o fim destes sons nasais. No primeiro caso analisou-se a posição da língua na articulação das vogais nasais (e orais para comparação) comparando os resultados para as seis regiões de Portugal Continental. Foram considerados os dados de todos os falantes das seis regiões contempladas, e, de modo a evitar o efeito do género do informante, as análises foram efectuadas, sempre em separado, para os dados de informantes do sexo masculino e feminino. Tal como no resto do trabalho, apenas se efectuaram análises para vogais nasais que ocorrem entre oclusivas (não nasais), para que o seu início tenha

as características de “vogal oral” tornando aplicável o processo de extracção e análise das formantes utilizado para as vogais orais. No que diz respeito à análise da fase final das vogais nasais foi estudada a dependência das características desta fase com a co-articulação com a consoante seguinte. As características espectrais no final das vogais nasais foram analisadas e compararam-se com as características espectrais das consoantes nasais e de outras vogais nasais. A semelhança espectral foi avaliada através de medidas que traduzem uma distância espectral.

Ao longo do trabalho foram, sempre que possível, efectuadas analogias entre os resultados alcançados e os obtidos anteriormente sobre este assunto, de modo a enriquecer a caracterização efectuada.

7.2 Resultados Importantes

O trabalho apresentado neste documento explora assuntos pouco analisados em estudos anteriores, que pretendem contribuir para uma melhor caracterização das vogais nasais do PE. Os resultados apresentados surgem, quer de evoluções de trabalhos efectuados anteriormente, como é exemplo o estudo apresentado no Capítulo 4 que foca a caracterização da fonte glotal na produção dos sons nasais, quer de análises novas. Como exemplo destas últimas, refere-se o estudo com o objectivo de obter informação sobre os segmentos relativos aos eventos de abertura nasal e fecho oral, de uma forma automática (secção 5.2.2).

Estudos recentes efectuados sobre as vogais nasais caracterizam a nasal como algo “dinâmico”, que inicia com uma configuração oral e segue para uma configuração “faringo-nasal”. É também para essa proposta que o trabalho apresentado neste documento tende.

Por último, é de notar que alguns dos estudos apresentados tem como objectivo principal contribuir para o esclarecimento de algumas controvérsias existentes, como é exemplo o estudo efectuado com o objectivo de analisar a relação das características da fase final das nasais com as consoantes nasais e outras vogais.

7.3 Trabalho Futuro

Na secção 5.2.2 desta dissertação foi apresentado um estudo exploratório com o objectivo de obter informação sobre as várias fases das vogais nasais de uma forma automática. Os resultados deste estudo, embora preliminares, sugerem a importância da existência de estudos mais detalhados sobre este tipo de segmentação automática. É, portanto, este um assunto que importa analisar num âmbito posterior a este trabalho.

O último capítulo deste documento diz respeito ao estudo com o objectivo de caracterizar o tracto durante a produção das nasais e nas diferentes fases que as constituem. Assim, as análises desenvolvidas no âmbito deste estudo contribuem para a caracterização, quer da fase inicial, quer da fase final das vogais nasais. O limite temporal inerente a este tipo de trabalhos

não permitiu o desenvolvimento do estudo com o objectivo de caracterizar a fase medial da vogal nasal. Assim, uma análise que recorra a métodos de processamento de sinal adequados à existência de pólos e zeros na função de transferência do tracto vocal (como, por exemplo, os apresentados por Karl e Arild (2002); Lindquist (2000); Byrnes *et al.* (2000)), de modo a caracterizar, pelo menos, a parte medial da realização deste tipo de sons, é, sem dúvida, um trabalho a efectuar futuramente, com um elevado grau de interesse.

Um outro estudo desenvolvido neste trabalho e que originou resultados pouco esclarecedores, diz respeito à análise da parte final da vogal nasal. Este análise pretendia esclarecer a existência, ou não, de relação de dependência entre a fase final das vogais nasais e a consoante seguinte. O trabalho baseou-se na aplicação de medidas que traduzem uma distância espectral, de modo a aferir a semelhança entre os segmentos relevantes, e acabou por se revelar inconclusivo pelo facto de apresentar resultados contraditórios. A aplicação de medidas de distância que tenham em consideração o mecanismo de percepção humano (De, 1993), é sugerida, então como trabalho a realizar posteriormente.

Para finalizar esta contribuição, e tendo em consideração os objectivos que foram propostos - a melhoria do conhecimento dos sons nasais, em particular as vogais, para a aplicação à síntese articulatória - sugere-se a utilização dos resultados apresentados ao longo do trabalho, na produção dos sons analisados, com o sintetizador articulatório em desenvolvimento na Universidade de Aveiro.

Bibliografia

- A. ALMEIDA (1976). The portuguese nasal vowels: Phonetics and phonemics. In J. SCHMIDT-RADEFELT (editor), *Readings in Portuguese Linguistics*, páginas 348–396. North Holland, New York.
- B. ATAL, J. L. MILLER, E R. D. KENT (editores) (1991). *Papers in Speech Communication: Speech Processing*. Acoustical Society of America. Ver também Kent *et al.* (1991); Miller *et al.* (1991).
- A. TEIXEIRA, L. FERREIRA, L. MOUTINHO, R. L. COIMBRA, E R. LISBOA (2004a). An acoustic corpus contemplating regional variation for studies of european portuguese nasals. In *Proc. 2004 Internacional Conference Language Resources and Evaluation*. Lisboa, Portugal.
- A. TEIXEIRA, F. VAZ, L. MOUTINHO, R. COIMBRA, E R. LISBOA (2004b). Para a melhoria da síntese articulatória das vogais nasais do português europeu. *Revista de Estudos da Linguagem (The Journal of Language Studies)*.
- R. J. BAKEN E R. F. ORLIKOFF (2000). *Clinical Measurements of Speech and Voice*. Singular Thompson Learning.
- J. M. BARBOSA (1961). Les voyelles nasales portugaises: Interpretation phonologique. In A. SOVIJÄRVI E P. AALTO (editores), *Proc. Fourth International Congress of Phonetic Sciences*, páginas 691–709. Mouton & Co, The Hague.
- M. BÅVEGÅRD, G. FANT, J. GAUFFIN, E J. LILJENCRAANTS (1993). Vocal tract sweeptone data and model simulations of vowels, laterals and nasals. *Speech Transmission Laboratory, Quarterly Progress and Status Report*, STL-QPSR 4:43–75.
- P. S. BEDDOR (1983). *Phonologic and Phonetic Effects of Nasalization on Vowel Height*. Indiana University Linguistics Club.
- J. S. BRIDLE E N. C. SEDGWICK (1977). A method for segmenting acoustic patterns, with applications to automatic speech recognition. In *Proc. ICASSP*, volume 2, páginas 656–659.
- C. I. BYRNES, T. T. GEORGIU, E A. LINDQUIST (2000). A new approach to spectral estimation: A tunable high-resolution spectral estimator. In *IEEE Trans. on Signal Processing*, volume 48, páginas 3189–3205.
- D. G. CHILDERS (2000). *Speech Processing and Synthesis Toolboxes*. John Wiley & Sons.

- A. DE (1993). *Auditory Distortion Measures for Speech Coder Evaluation*. Tese de Doutorado, McGill University, Montréal, Canada.
- E. M. G. DE SOUSA (1994). *Para a Caracterização Fonético-Acústica da Nasalidade no Português do Brasil*. Dissertação de mestrado em linguística, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Brasil.
- P. DELATTRE (1954). Les attributs acoustiques de la nasalité vocalique et consonantique. *Studia Linguist.*, 8:103–109.
- M. R. DELGADO-MARTINS (2002). *Fonética do Português. Trinta Anos de Investigação*. Editora Caminho, Lisboa.
- J. DELLER, J. G. PROAKIS, E J. H. L. HANSEN (1993). *Discrete Time Processing of Speech Signals*. MacMillan, NY.
- M. DRENSKA (1988). Análise acústica das vogais nasais em português e búlgaro. In *Actas do 4o. Encontro da Associação Portuguesa de Linguística*, páginas 139–165.
- G. FANT (1960). *Acoustic theory of speech production*. Mouton and Co., Gravenhage, The Netherlands. Second Edition in 1970.
- O. FUJIMURA (1960). Spectra of nasalized vowels. *Quarterly Progress Report, Research Laboratory of Technology, M. I. T.*, 58:214–218.
- M. J. GALVÃO (1998). The nasal vowels in iberian portuguese. In P. K. KUHL E L. A. CRUM (editores), *16th International Congress on Acoustics and the 135th Meeting of the Acoustical Society of America*, páginas 2949–2950. Seattle, Washington, USA.
URL <http://www.apl.washington.edu/asa/asa.html>
- A. H. GRAY JR. E J. D. MARKEL (1976). Distance measures for speech processing. *IEEE Trans. Acoustics, Speech and Signal Proc.*, ASSP-24(5):380–391. Reimpresso em Atal *et al.* (1991).
- R. C. GUIMARÃES E J. A. S. CABRAL (1997). *Estatística*. McGraw Hill, Amadora.
- S. HAWKINS E K. N. STEVENS (1985). Acoustic and perceptual correlates of the non-nasal–nasal distinction for vowels. *Journal of the Acoustical Society of America*, 77(4):1560–1575.
- H. HERMANSKY (1990). Perceptual linear predictive (PLP) analysis of speech. *Journal of the Acoustical Society of America*, 87(4):1738–1752.
- H. HERMANSKY E N. MORGAN (1994). RASTA processing of speech. *IEEE Trans. on Speech Audio Processing*, 2(4):578–589.
- A. S. HOUSE E K. S. STEVENS (1956). Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders*, 21(2):218–232.

- M. K. HUFFMAN E R. A. KRAKOW (editores) (1993). *Nasals, Nasalization, and the Velum*. Phonetics and Phonology (vol. 5). Academic Press Inc.
- S. KARL E L. ARILD (2002). Analysis of nasals and nasalized vowels based on branched tube models. In *European Signal Processing Conference*, volume 3, páginas 65–68.
- H. KAWASAKI (1986). Phonetic explanation for phonological universals: the case of distinctive vowel nasalization. In J. J. OHALA E J. J. JAEGER (editores), *Experimental Phonology*, páginas 81–103. Academic Press, New York.
- R. D. KENT, B. ATAL, E J. L. MILLER (editores) (1991). *Papers in Speech Communication: Speech Production*. Acoustical Society of America. Ver também Atal *et al.* (1991); Miller *et al.* (1991).
- R. A. KRAKOW, P. S. BEDDOR, L. M. GOLDSTEIN, E C. FOWLER (1988). Coarticulatory influences on the perceived height of nasal vowels. *Journal of the Acoustical Society of America*, 83(3):1146–1158.
- A. LACERDA E B. F. HEAD (1966). Análise de sons nasais e sons nasalizados do português. *Revista do Laboratório de Fonética Experimental de Coimbra*, VI:5–70.
- A. LINDQUIST (2000). Some new methods and concepts in high-resolution spectral estimation. In *Plenary Lecture Third Asian Control Conference*. Shangai.
- J. N. LITTLE E L. SHURE (1992). *Signal Processing TOOLBOX - For use with Matlab*. The MathWorks, Inc. User's Guide.
- R. H. MANNELL (1994). The prediction of “perceptual distance” from spectral distance measures based upon auditory and non-auditory models of intensity scaling. In *Fifth Australian International Conference on Speech Science and Technology*. Perth.
- M. H. M. MATEUS (1975). *Aspectos de Fonologia Portuguesa*. Publicações do Centro de Estudos Filológicos, Lisboa.
- J. L. MILLER, R. D. KENT, E B. S. ATAL (editores) (1991). *Papers in Speech Communication: Speech Perception*. Acoustical Society of America. Ver também Kent *et al.* (1991); Atal *et al.* (1991).
- L. C. MOUTINHO (2000). *Uma introdução ao Estudo da Fonética e Fonologia do Português*. Plátano Edições Técnicas, primeira edição.
- L. RABINER E B.-H. JUANG (1993). *Fundamentals of Speech Recognition*. Signal Processing Series, Alan V. Oppenheim, Series Editor. Prentice Hall.
- L. R. RABINER, M. J. CHENG, A. E. ROSENBERG, E C. A. MCGONEGAL (1976). A comparative performance study of several pitch detection algorithms. *IEEE Trans. Acoustics, Speech and Signal Proc.*, ASSP-24(5):399–418.

- R. SAMPSON (editor) (1999). *Nasal Vowel Evolution in Romance*. Oxford University Press. ISBN 0-19823848-7.
- B. G. SECREST E G. R. DODDINGTON (1983). An integrated pitch tracking algorithm for speech systems. *IEEE Trans. Acoustics, Speech and Signal Proc.*, 8:1352–1355.
- A. R. T. V. SILVA (1995). *Análise de Fonemas Nasais da Língua Portuguesa*. Dissertação de mestrado em electrónica e telecomunicações, Departamento de Electrónica e Telecomunicações, Universidade de Aveiro, Portugal.
- K. N. STEVENS, A. ANDRADE, E M. C. VIANA (1987). Perception of vowel nasalization in VC contexts: A cross-language study. *Journal of the Acoustical Society of America*, 82:S119.
- K. N. STEVENS, S. KASOWSKI, E C. G. FANT (1953). An electrical analog of the vocal tract. *Journal of the Acoustical Society of America*, 25:734–742.
- S. TAKEUCHI, H. KASUYA, E K. KIDO (1975). On the acoustic correlate of nasality. *Journal Acoustical Society of Japan*, 31(5):298–309.
- A. TEIXEIRA (2000). *Síntese Articulatoria das Vogais Nasais do Português Europeu*. Tese de doutoramento, Universidade de Aveiro.
- A. TEIXEIRA, L. FERREIRA, R. L. COIMBRA, E L. C. MOUTINHO (2003a). Estudo da posição da língua no início das vogais nasais de seis regiões de Portugal pela análise das duas primeiras formantes. Relatório técnico, Departamento de Línguas e Culturas e Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA).
- A. TEIXEIRA, L. MOUTINHO, E R. COIMBRA (2003b). Production, acoustic and perceptual studies on European Portuguese nasal vowels height. In *International Congress of Phonetic Sciences*, páginas 3033–3036. Barcelona.
- A. TEIXEIRA E F. VAZ (2001). European Portuguese nasal vowels: An EMMA study. In *Proc. EuroSpeech*.
- A. R. G. VIANA (1883). *Essai de phonétique et de phonologie de la langue portugaise d'après le dialecte actuel de Lisbonne*. Paris.

Apêndice A

Constituição do *corpus* Regional

t01i01	pantufa
t01i02	tampa
t01i03	estante
t01i04	tanque
t01i05	campo
t01i06	pente
t01i07	penca
t01i08	oitenta
t01i09	quente
t01i10	pintor
t01i11	tinto
t01i12	quinta
t01i13	ponte
t01i14	compras
t01i15	conta
t01i16	conquistador
t01i17	cumprimentar

Tabela A.1: Vogais Nasais entre oclusivas surdas.

t02i01	bambu
t02i02	bandeira
t02i03	gamba
t02i04	propaganda
t02i05	ganga
t02i06	cabendo
t02i07	bengala
t02i08	podendo
t02i09	bimbo
t02i10	bingo
t02i11	guindaste
t02i12	bombeiro
t02i13	redondo
t02i14	candongueiro
t02i15	gôndola
t02i16	gongo
t02i17	bumba
t02i18	segunda

Tabela A.2: Vogais Nasais entre oclusivas sonoras.

t03i01	pandeireta
t03i02	tambor
t03i03	tango
t03i04	cambalhota
t03i05	candeeiro
t03i06	canguru
t03i07	setembro
t03i08	estendal
t03i09	pimba
t03i10	pinguim
t03i11	pombo
t03i12	comboio
t03i13	condão
t03i14	rotunda
t03i15	secundário

Tabela A.3: Vogais Nasais após oclusivas surdas e antes de oclusivas sonoras.

t04i01	banco
t04i02	estudante
t04i03	gigante
t04i04	rebentar
t04i05	dente
t04i06	pedinte
t04i07	seguinte
t04i08	mastodonte
t04i09	pergunta

Tabela A.4: Vogais Nasais após oclusivas sonoras e antes de oclusivas surdas.

t05i01	fanfarra
t05i02	chanfrado
t05i03	avançar
t05i04	ascensor
t05i05	gengiva
t05i06	vencimento
t05i07	ginjinha
t05i08	cinzento
t05i09	província
t05i10	afonso
t05i11	sonso
t05i12	zonzó
t05i13	funchal
t05i14	jejum
t05i15	fungicida

Tabela A.5: Vogais Nasais entre fricativas.

t06i01	amamos
t06i02	amámos
t06i03	manta
t06i04	mancha
t06i05	membrana
t06i06	mentiroso
t06i07	mensal
t06i08	domingo
t06i09	península
t06i10	ninfa
t06i11	conosco
t06i12	monte
t06i13	monge
t06i14	mundo
t06i15	anúncio

Tabela A.6: Palavras com Consoantes Nasais.

t07i01	champô
t07i02	sandes
t07i03	jantar
t07i04	penso
t07i05	agenda
t07i06	benfica
t07i07	gente
t07i08	pincel
t07i09	cinto
t07i10	vindima
t07i11	vinte
t07i12	esponja
t07i13	sombra
t07i14	assunto
t07i15	fundo

Tabela A.7: Vogais Nasais antes ou após uma oclusiva.

t08i01	lâmpada
t08i02	rampa
t08i03	lenço
t08i04	renda
t08i05	cilindro
t08i06	limpar
t08i07	lomba
t08i08	roncar
t08i09	lumbago

Tabela A.8: Vogais Nasais após consoante líquida.

t09i01	pato
t09i02	pacote
t09i03	tapete
t09i04	capacete
t09i05	capacete
t09i06	baba
t09i07	gaguejar
t09i08	peca
t09i09	pateta
t09i10	degrau
t09i11	tipo
t09i12	tito
t09i13	guida
t09i14	hipopótamo
t09i15	pote
t09i16	totoloto
t09i17	copo
t09i18	bigode
t09i19	cupido
t09i20	buda
t09i21	narigudo
t09i22	pintor
t09i23	conquistador
t09i24	tambor
t09i25	mentiroso
t09i26	connosco
t09i27	totoloto
t09i28	tapete
t09i29	capacete
t09i30	ampulheta

Tabela A.9: Vogais Orais em vários contextos.

t10i01	ampulheta
t10i02	lã
t10i03	encontro
t10i04	entrada
t10i05	importante
t10i06	rim
t10i07	onze
t10i08	umbigo
t10i09	atum

Tabela A.10: Vogais Nasais em início ou fim de palavra.

t11i01	lã azul
t11i02	tom adequado
t11i03	fim anunciado
t11i04	fim esperado
t11i05	com manteiga
t11i06	com naturalidade
t11i07	tom natural
t11i08	lã antiga
t11i09	som ambiente
t11i10	manhã anterior
t11i11	a antiguidade

Tabela A.11: Sequências contendo vogais nasais nos limites das palavras.

t12i01	mãe
t12i02	também
t12i03	põe
t12i04	muito
t12i05	coincidente
t12i06	quando
t12i07	fiando

Tabela A.12: Ditongos Nasais.

Apêndice B

Características dos Informantes

Nome	Idade	Sexo	Nível de Escolaridade	Local de Origem	Região
ALT	74	F	freq. 3a classe	Vila Viçosa	Alentejo
JQM	42	M	freq. 5o ano	S. Miguel de Machede	Alentejo
JSD	78	M	não sabe ler	Vila Viçosa	Alentejo
ARS	42	F	3a classe	S. Miguel de Machede	Alentejo
NMT	76	F	4a classe	S. Bartolomeu de Messines	Algarve
ACR	62	F	3a classe	S. Bartolomeu de Messines	Algarve
JLM	69	M	4a classe	Vale de Parra (Albufeira)	Algarve
APM	81	M	só sabe assinar o nome	S. Bartolomeu de Messines	Algarve
AMT	76	M	Escola Primária	S. Bartolomeu de Messines	Algarve
RAL	74	F	3a classe	S. Pedro do Sul	Beira Interior
JRD	82	M	3a classe / 4a classe em adulto	S. Pedro do Sul	Beira Interior
CCD	62	F	3a classe / 4a classe em adulto	S. Pedro do Sul	Beira Interior
JLZ	64	M	6o ano	S. Pedro do Sul	Beira Interior
PRQ	33	M	9o ano inacabado	Aveiro	Beira Litoral
CLR	52	M	6o ano	Aveiro	Beira Litoral
OCN	40	F	6o ano	Aveiro	Beira Litoral
NSB	52	F	6o ano	Aveiro	Beira Litoral
ALS	61	M	4a classe	Sequeiros	Minho
MMC	44	F	6o ano	Sequeiros	Minho
ALD	61	M	4a classe	Salamonde	Minho
ADL	55	F	4a classe	Salamonde	Minho
APC	35	M	7o ano	Podence	Trás-os- Montes
DRO	81	M	4a classe	Podence	Trás-os-Montes
CVA	74	F	4a classe inacabada	Podence	Trás-os-Montes
SCF	77	F	4a classe	Podence	Trás-os-Montes

Tabela B.1: Informantes do *corpus* Regional.

Apêndice C

Estímulos Visuais



Figura C.1: Imagem correspondente à palavra *Candeeiro*.



Figura C.2: Imagem correspondente à palavra *Tambor*.

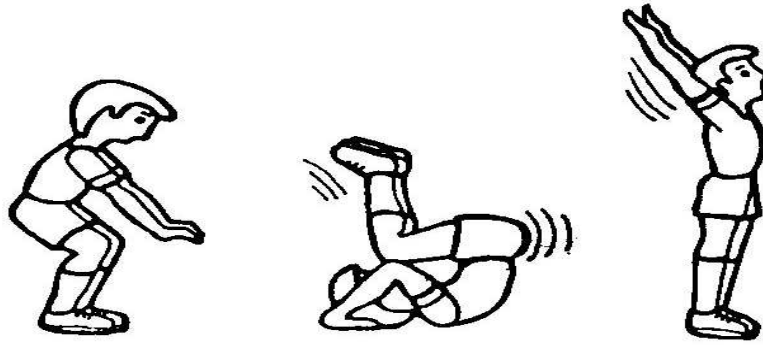


Figura C.3: Imagem correspondente à palavra *Cambalhota*.



Figura C.4: Imagem correspondente à palavra *Estendal*.



Figura C.5: Imagem correspondente à palavra *Pombo*.



Figura C.6: Imagem correspondente à palavra *Comboio*.



Figura C.7: Imagem correspondente à palavra *Pandeireta*.



Figura C.8: Imagem correspondente à palavra *Pinguim*.

Apêndice D

Resultados Características da Fonte sem médias

D.1 Frequência Fundamental

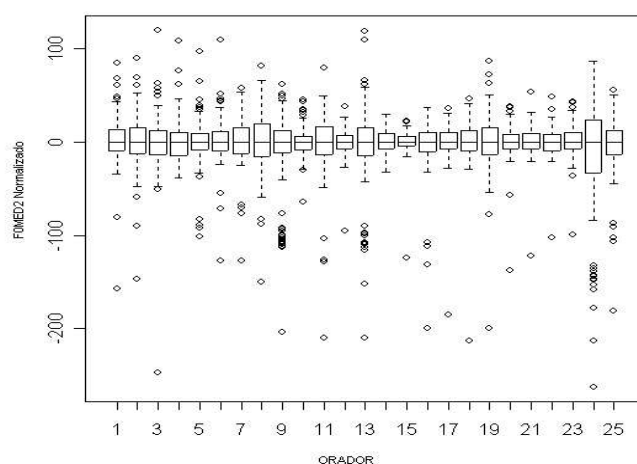


Figura D.1: Diagrama de extremos e quartis dos valores de F0MED em função do falante após subtração da mediana de F0MED.

D.2 Quociente de Abertura

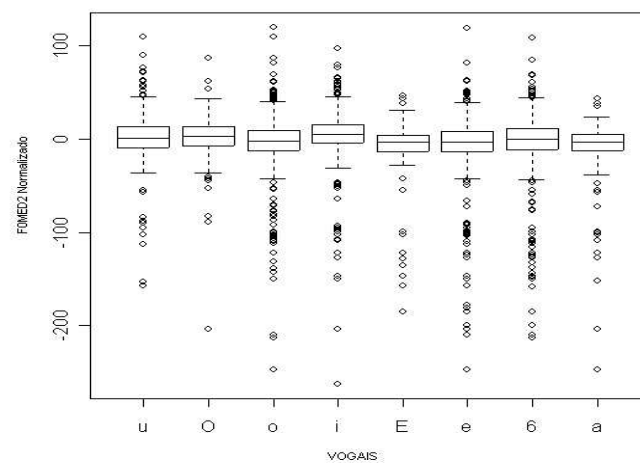


Figura D.2: Diagrama de extremos e quartis dos valores de F0MED em função da vogal após subtracção da mediana de F0MED.

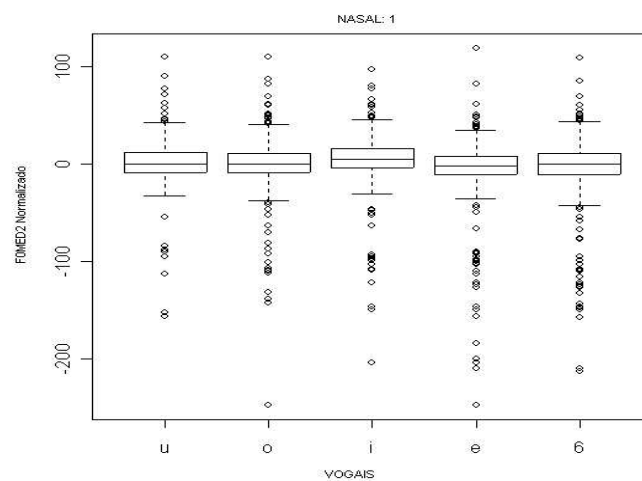


Figura D.3: Diagrama de extremos e quartis dos valores de F0MED em função das vogais nasais após subtracção da mediana de F0MED.

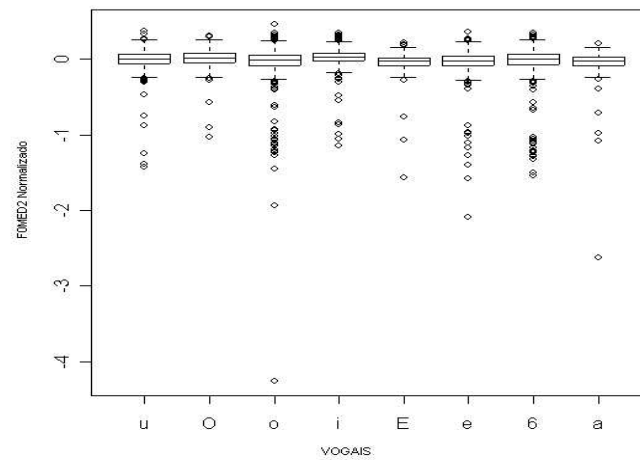


Figura D.4: Diagrama de extremos e quartis dos valores de F0MED em função da vogal após subtração da mediana de F0MED e divisão por F0MED.

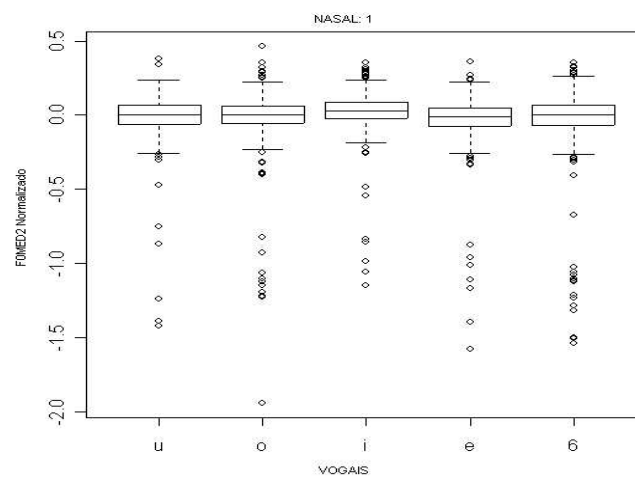


Figura D.5: Diagrama de extremos e quartis dos valores de F0MED em função das vogais nasais após subtração da mediana de F0MED e divisão por F0MED.

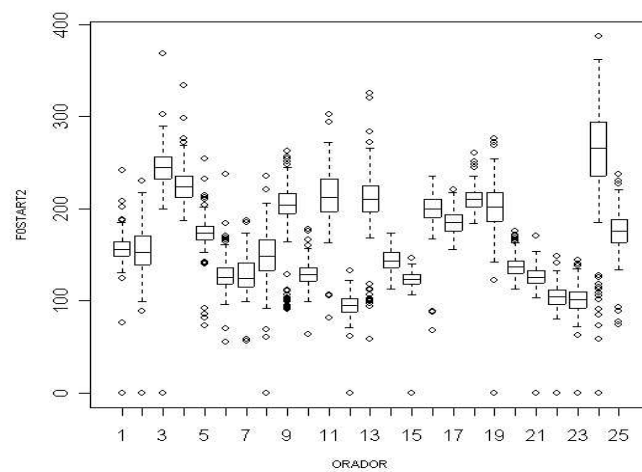


Figura D.6: Diagrama de extremos e quartis dos valores de F0START em função do falante.

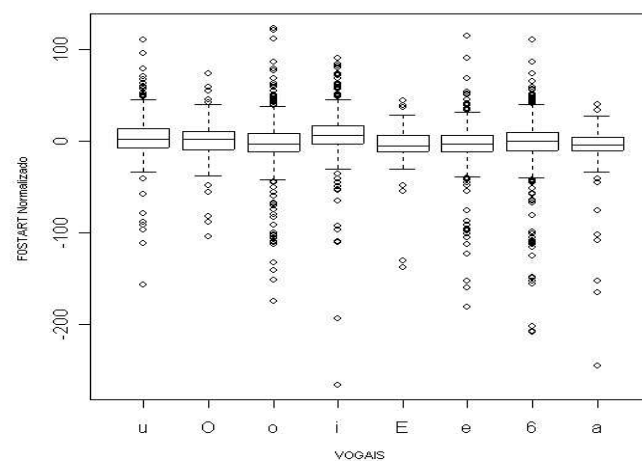


Figura D.7: Diagrama de extremos e quartis dos valores de F0START em função da vogal após subtração da mediana de F0START.

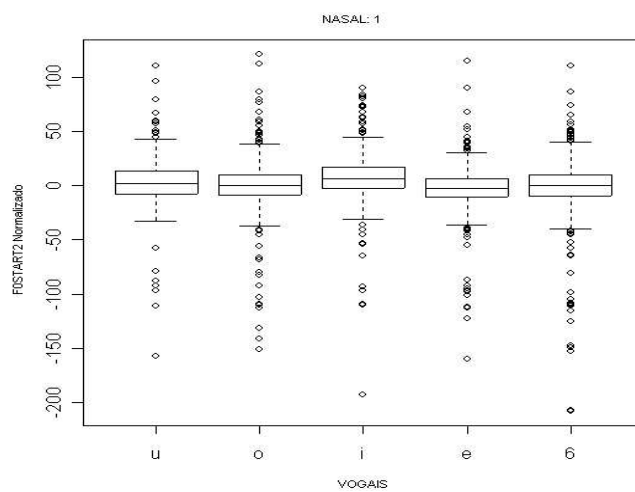


Figura D.8: Diagrama de extremos e quartis dos valores de F0START em função das vogais nasais após subtração da mediana de F0START.

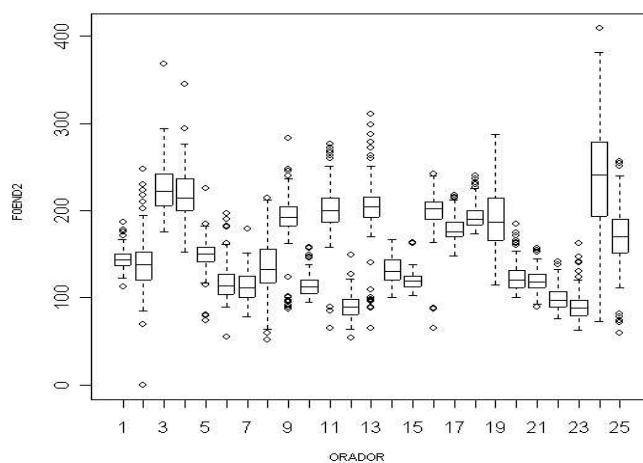


Figura D.9: Diagrama de extremos e quartis dos valores de F0END em função do falante.

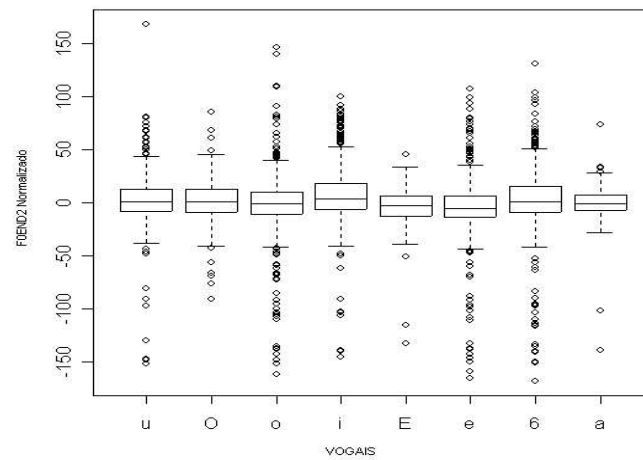


Figura D.10: Diagrama de extremos e quartis dos valores de F0END em função da vogal após subtração da mediana de F0END.

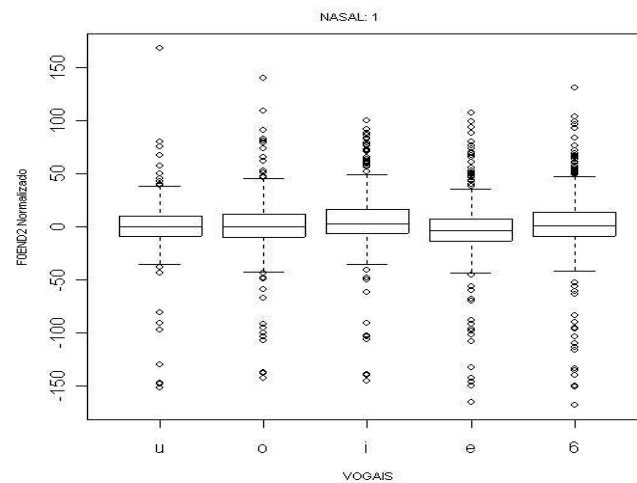


Figura D.11: Diagrama de extremos e quartis dos valores de F0END em função das vogais nasais após subtração da mediana de F0END.

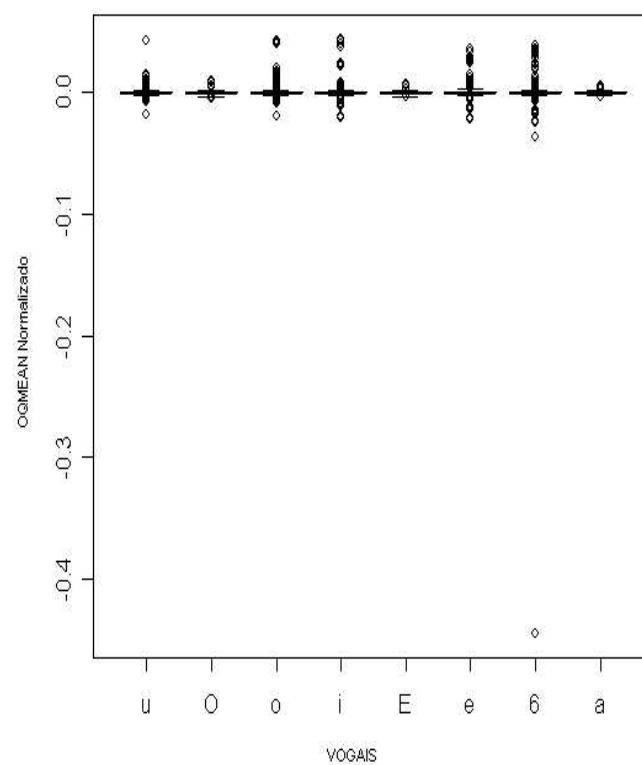


Figura D.12: Diagrama de extremos e quartis dos valores de OQMean em função da vogal após subtração da mediana de OQMEan.

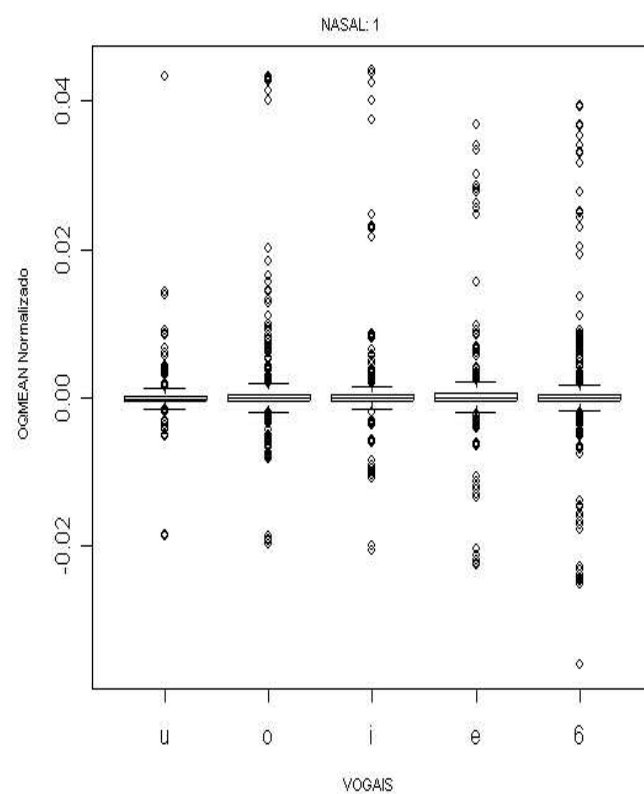


Figura D.13: Diagrama de extremos e quartis dos valores de OQMean em função das vogais nasais após subtração da mediana de OQMean.

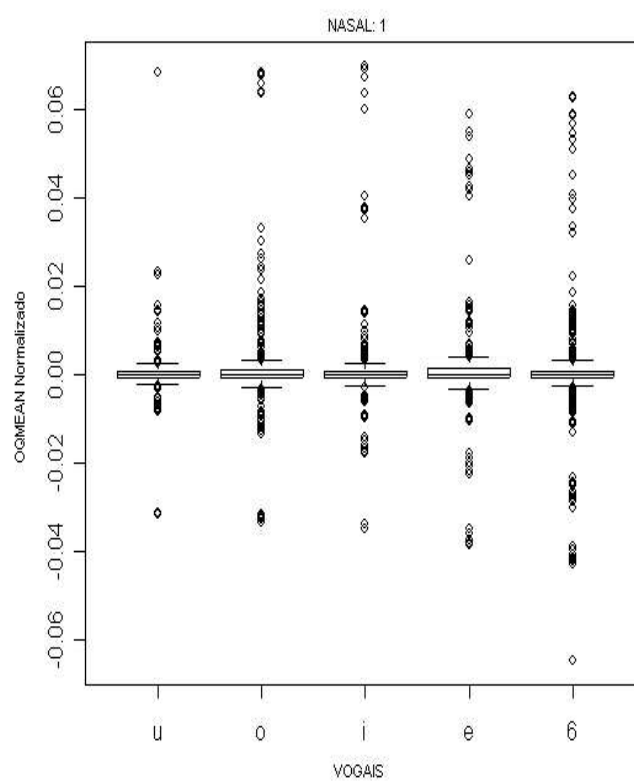


Figura D.14: Diagrama de extremos e quartis dos valores de OQMean em função das vogais nasais após subtração da mediana de OQMean e divisão por OQMean.

Apêndice E

Resultados da Caracterização das Fases das Vogais Nasais com Base no Sinal Acústico

NasalMid - Seg2

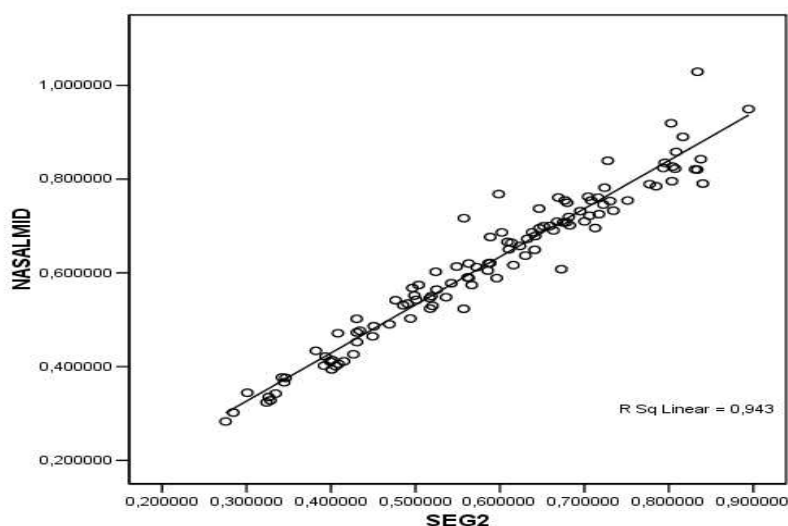


Figura E.1: *Scatter Plot* entre o evento médio do velo e o segundo segmento automático (em três).

Regressão

Modelo	R	R Square	Adjusted R Square
1	,971	,943	,943

Tabela E.1: Sumário do Modelo obtido para o evento médio do velo e o segundo segmento automático (em três).

A recta de regressão é:

$$NasalMid = 1,026 \times SEG2 - 0,019 \quad (E.1)$$

NasalClose - Seg3

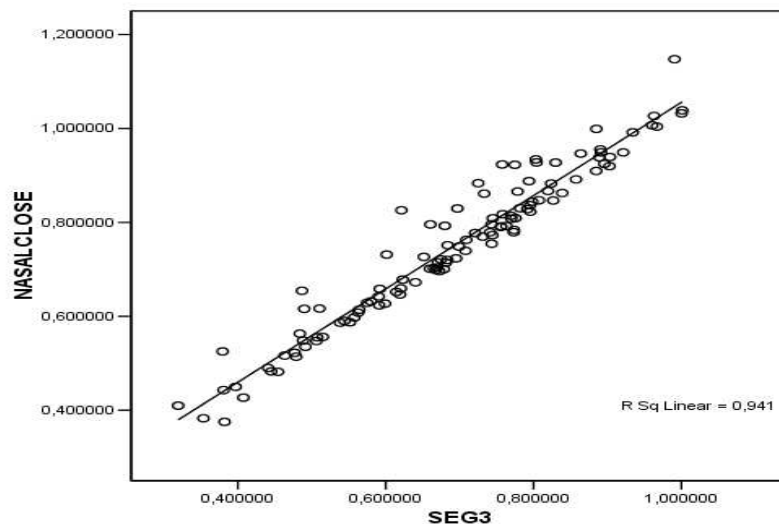


Figura E.2: *Scatter Plot* entre o evento de fecho do velo e o terceiro segmento automático (em três).

Regressão

Modelo	R	R Square	Adjusted R Square
1	,970	,941	,940

Tabela E.2: Sumário do Modelo obtido para o evento de fecho do velo e o terceiro segmento automático (em três).

A recta de regressão é:

$$NasalClose = 0,993 \times SEG3 - 0,063 \quad (E.2)$$

OralOpen - Seg2

Regressão

Modelo	R	R Square	Adjusted R Square
1	,855	,730	,728

Tabela E.3: Sumário do Modelo obtido para o evento de abertura dos articuladores orais e o segundo segmento automático (em três).

A recta de regressão é:

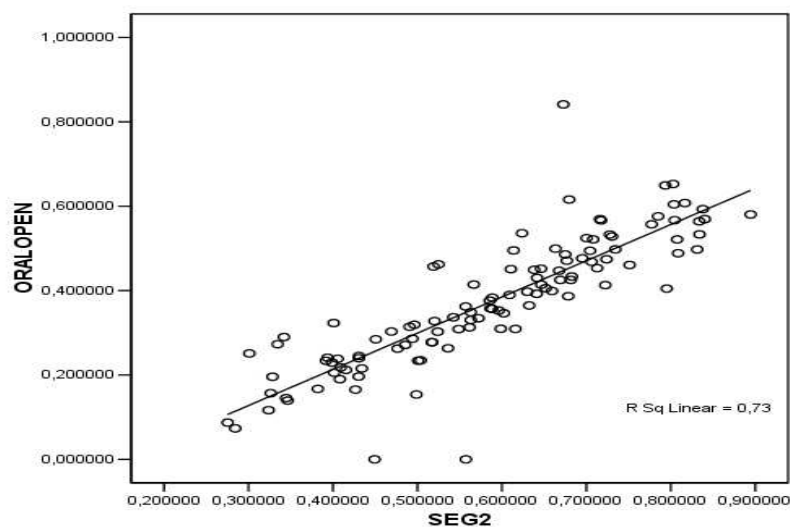


Figura E.3: *Scatter Plot* entre o evento de abertura dos articuladores orais e o segundo segmento automático (em três).

$$OralOpen = 0,858 \times SEG2 - 0,130 \quad (E.3)$$

NasalClose - Seg4

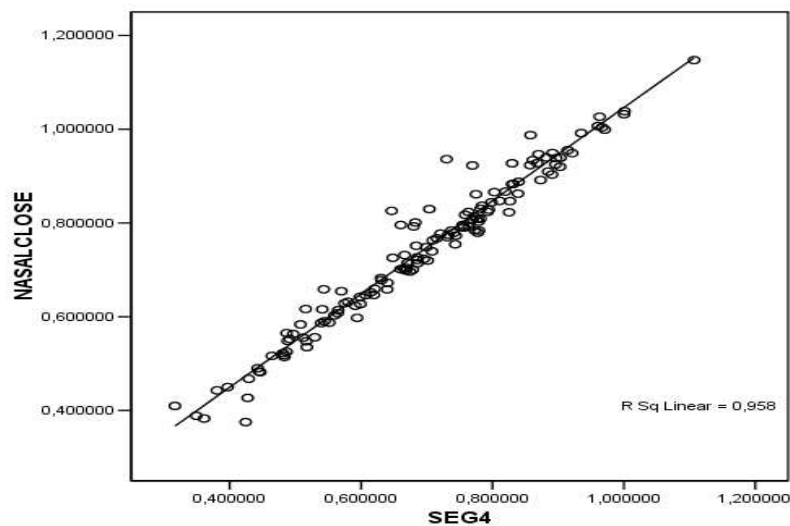


Figura E.4: *Scatter Plot* entre o evento de fecho do velo e o quarto segmento automático (em 4 segmentos).

Regressão

A recta de regressão é:

Modelo	R	R Square	Adjusted R Square
1	,979	,958	,958

Tabela E.4: Sumário do Modelo obtido para o evento de fecho do velo e o quarto segmento automático (em 4 segmentos).

$$NasalClose = 0,994 \times SEG4 - 0,052 \quad (E.4)$$

OralOpen - Seg1

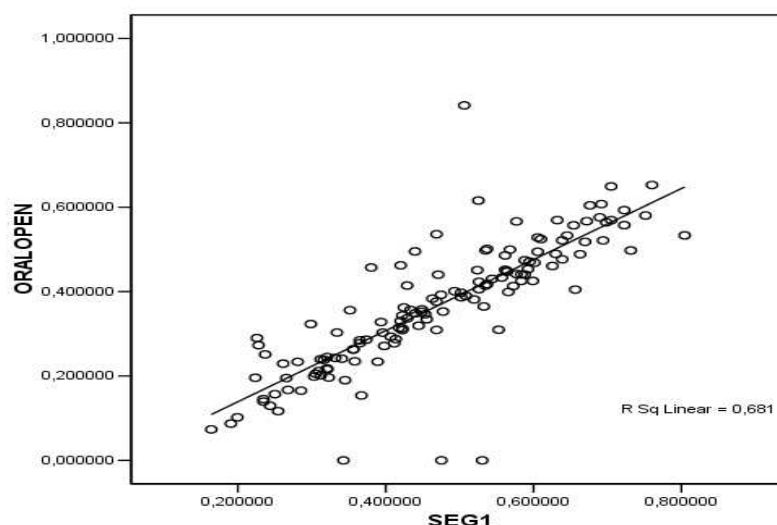


Figura E.5: *Scatter Plot* entre o evento de fecho dos articuladores orais e o primeiro segmento automático (em 4 segmentos).

Regressão

Modelo	R	R Square	Adjusted R Square
1	,825	,681	,679

Tabela E.5: Sumário do Modelo obtido para o evento de fecho dos articuladores orais e o primeiro segmento automático (em 4 segmentos).

A recta de regressão é:

$$OralOpen = 0,842 \times SEG1 - 0,029 \quad (E.5)$$

NasalMid - Seg2

Regressão

A recta de regressão é:

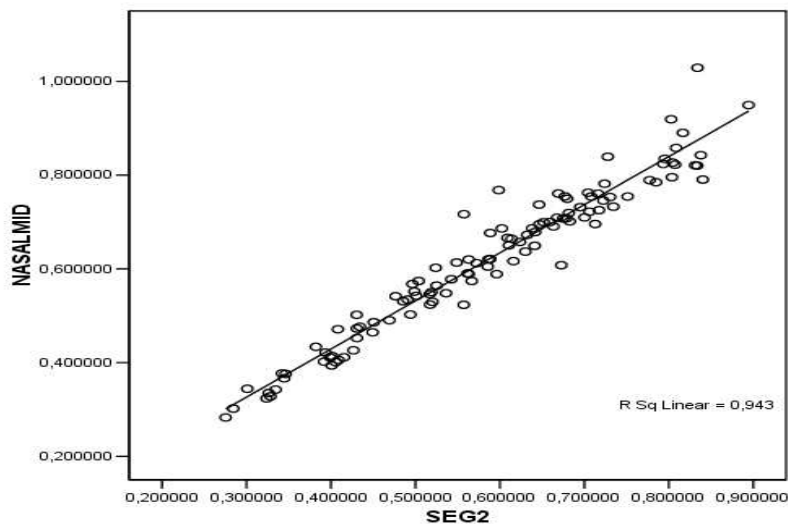


Figura E.6: *Scatter Plot* entre o evento médio do velo e o segundo segmento automático (em 4 segmentos).

Modelo	R	R Square	Adjusted R Square
1	,971	,943	,943

Tabela E.6: Sumário do Modelo obtido para o evento médio do velo e o segundo segmento automático (em 4 segmentos).

$$NasalMid = 1,026 \times SEG2 - 0,019 \quad (E.6)$$

NasalMid - Seg4

Regressão

Modelo	R	R Square	Adjusted R Square
1	,978	,956	,955

Tabela E.7: Sumário do Modelo obtido para o evento medial do velo e o quarto segmento automático (em 4 segmentos).

A recta de regressão é:

$$NasalMid = 0,961 \times SEG4 - 0,049 \quad (E.7)$$

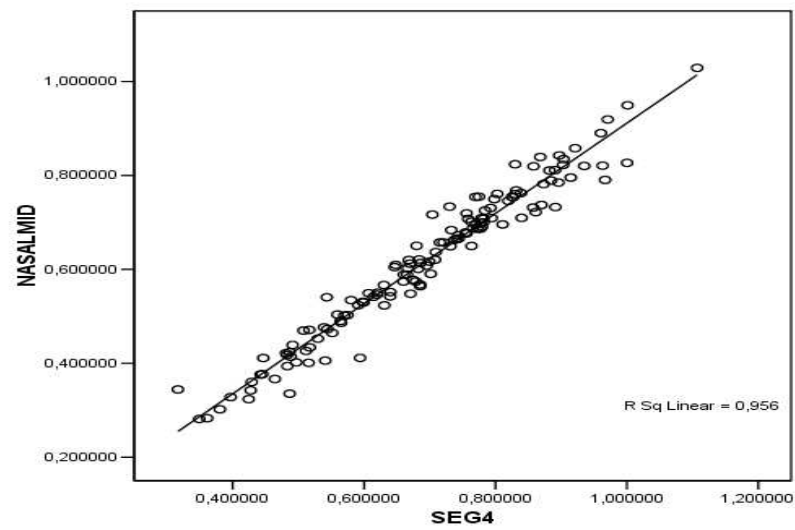


Figura E.7: *Scatter Plot* entre o evento medial do velo e o quarto segmento automático (em 4 segmentos).

Apêndice F

Resultados da Análise da Fase Inicial de uma Vogal Nasal

Análise Global de todos os dados

Os resultados apresentados nesta secção referem-se aos valores das duas primeiras formantes (F1 e F2) obtidos para cada uma das vogais em análise.

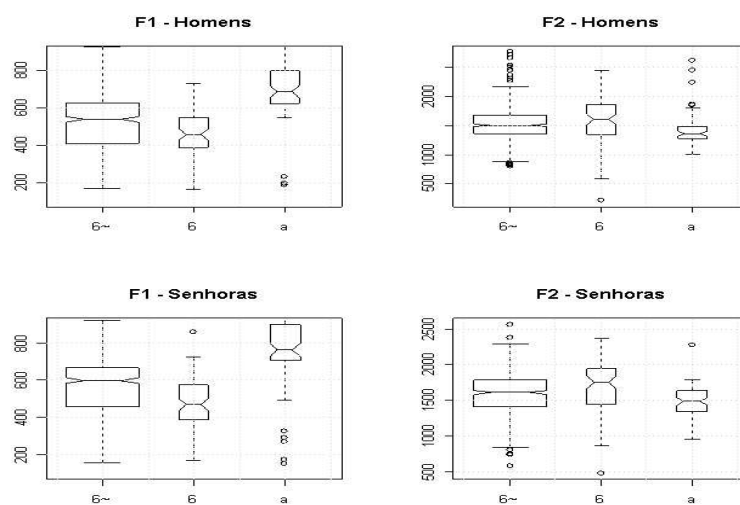


Figura F.1: Valores de F1 e F2 (à direita) para as vogais [6~][6] e [a].

Análise por Região

Apresenta-se também os diagramas de extremos e quartis para os valores de F1, F2 e para a diferença entre F2 e F1, para as várias regiões estudadas.

Assim, as figuras F.6, F.7, F.8 e F.9 apresentam os resultados obtidos para a região do MINHO.

Para a região de Trás-os-Montes são apresentadas as figuras F.10 a F.13.

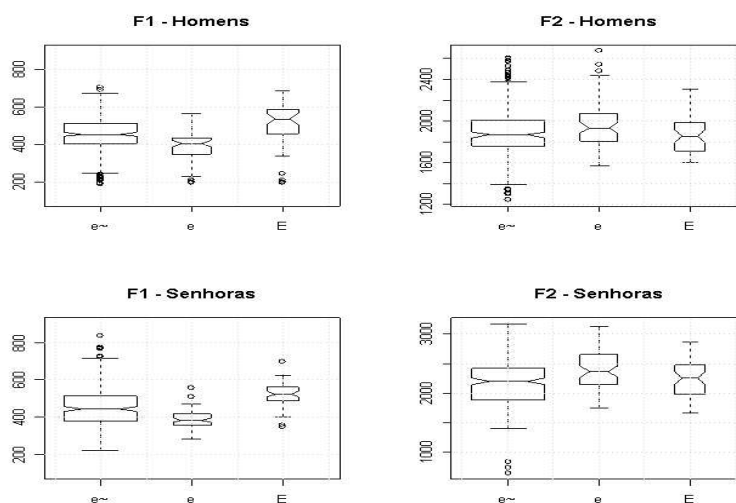


Figura F.2: Valores de F1 e F2 (à direita) para as vogais [e~][e] e [E].

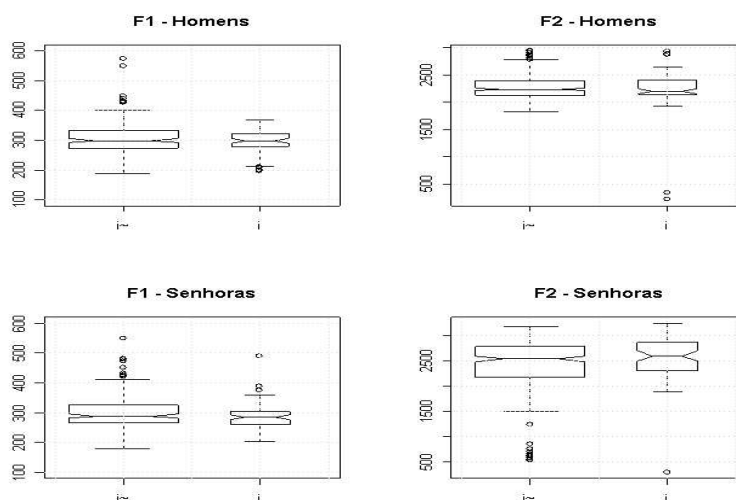


Figura F.3: Valores de F1 e F2 (à direita) para as vogais [i~] e [i].

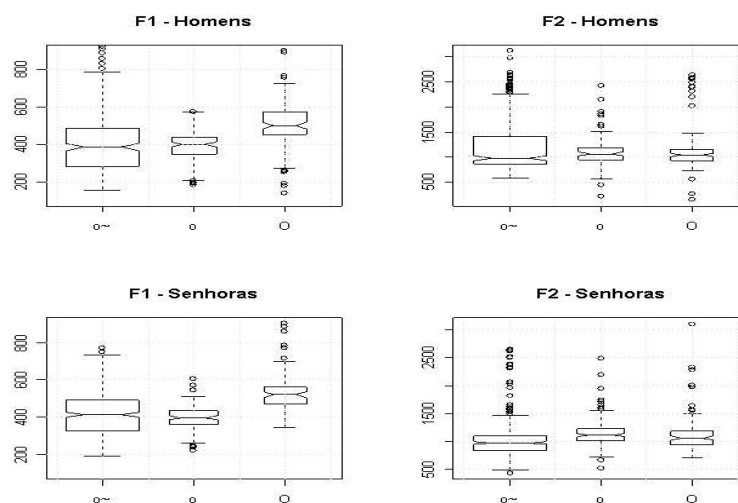


Figura F.4: Valores de F1 e F2 (à direita) para as vogais [õ][o] e [O].

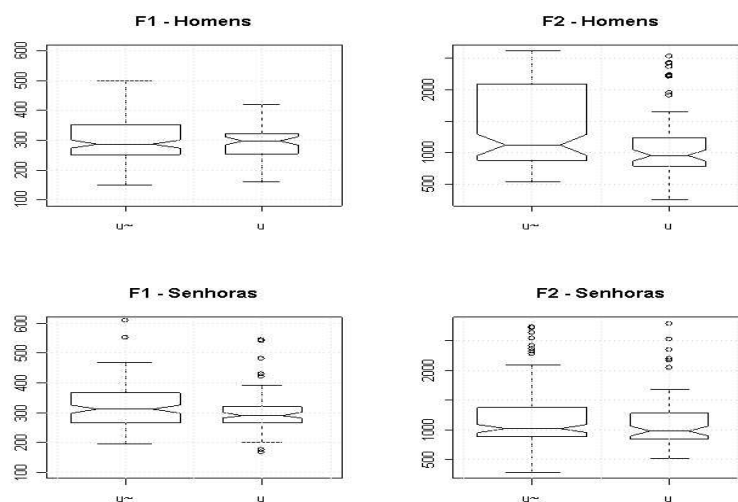


Figura F.5: Valores de F1 e F2 (à direita) para as vogais [ũ] e [u].

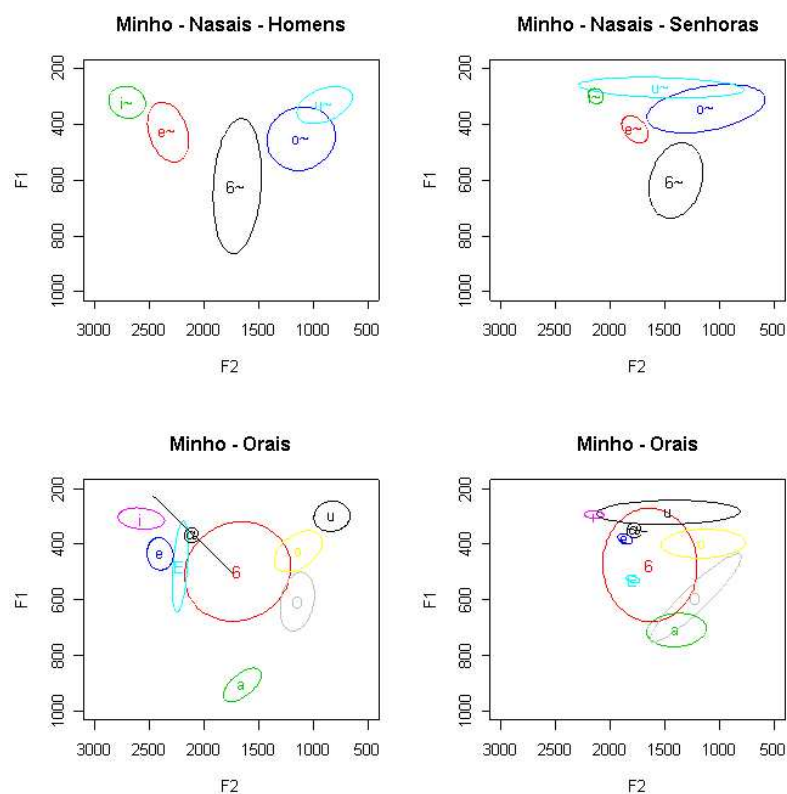


Figura F.6: Representação da distribuição das duas primeiras formantes para os falantes do MINHO.

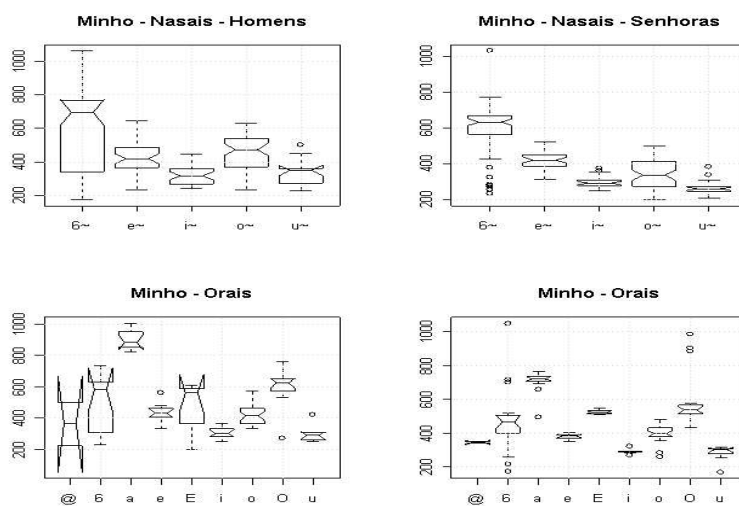


Figura F.7: Boxplots de F1 (separação por sexo e nasalidade) para os falantes do MINHO.

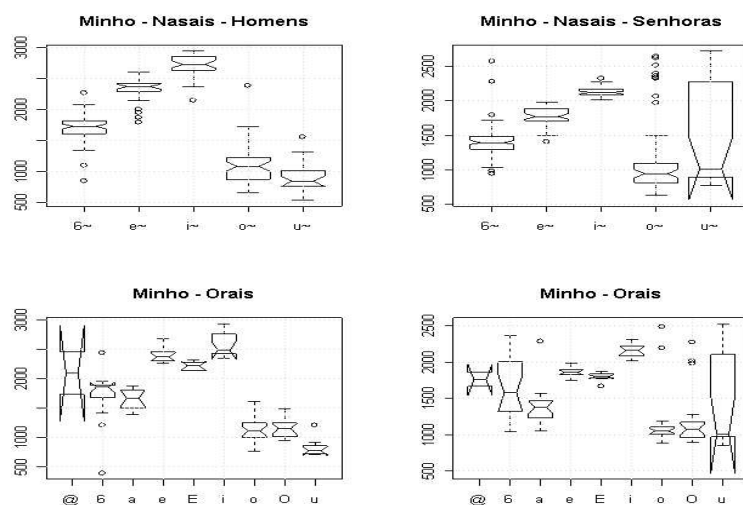


Figura F.8: Boxplots de F2 (separação por sexo e nasalidade) para os falantes do MINHO.

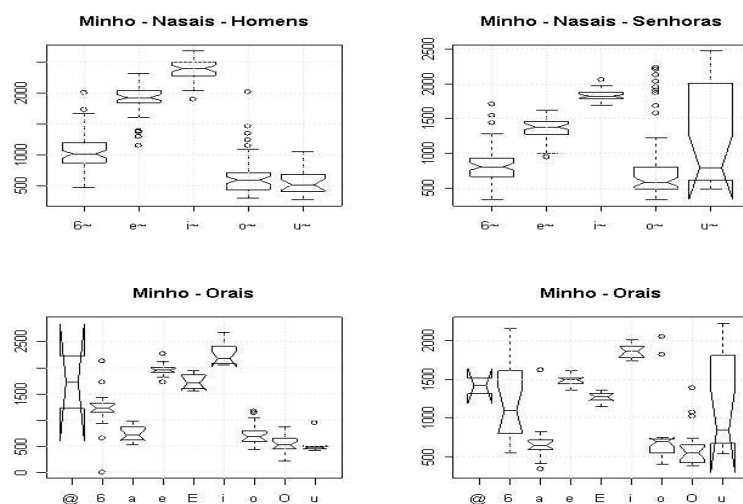


Figura F.9: Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes do MINHO.

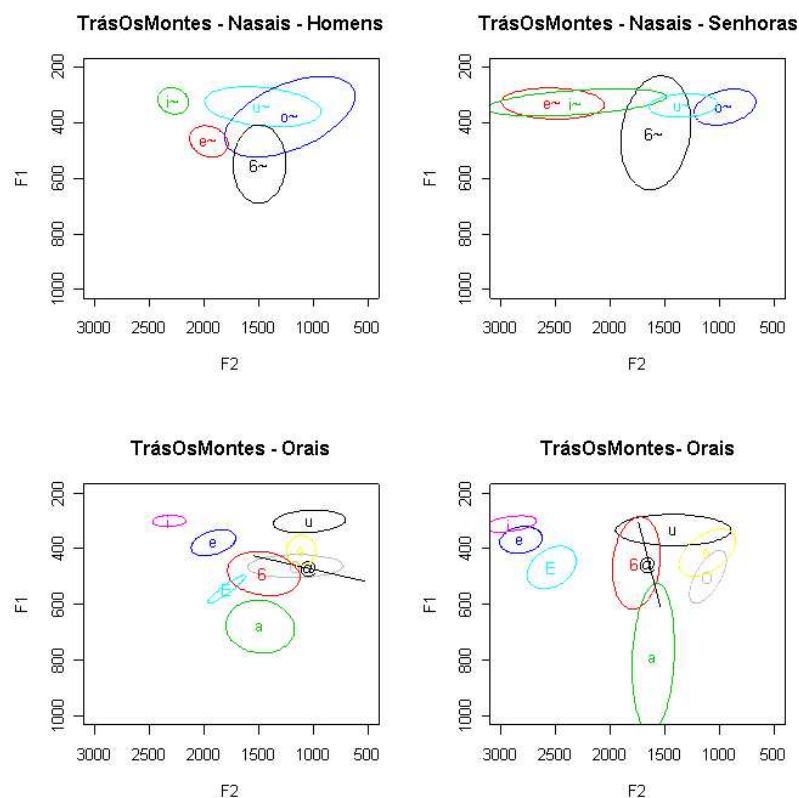


Figura F.10: Representação da distribuição das duas primeiras formantes para os falantes de TRÁS-OS-MONTES.

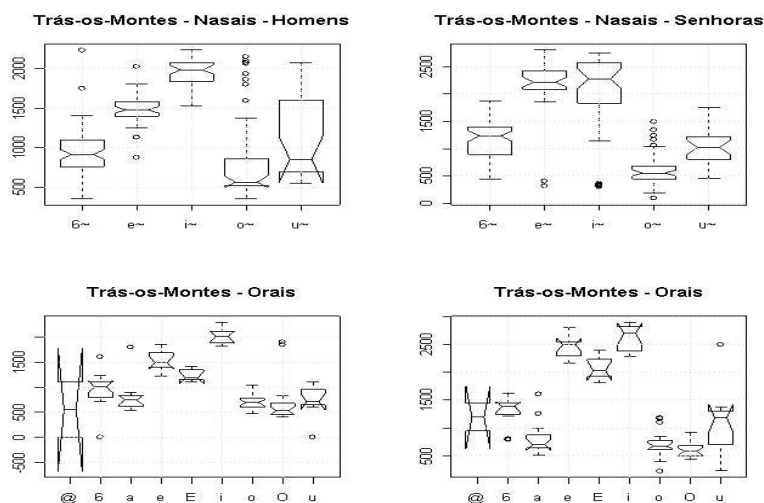


Figura F.11: Boxplots de F1 (separação por sexo e nasalidade) para os falantes de TRAS-OS-MONTES.

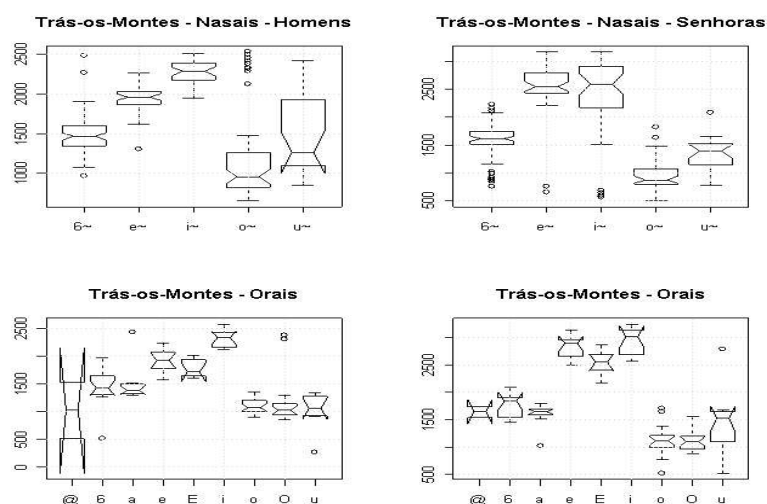


Figura F.12: Boxplots de F2 (separação por sexo e nasalidade) para os falantes de TRAS-OS-MONTES.

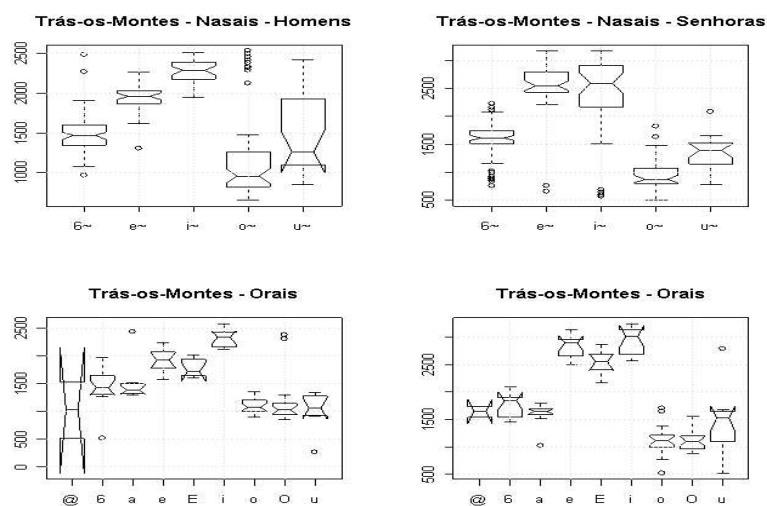


Figura F.13: Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes de TRAS-OS-MONTES.

Os resultados obtidos para a região da Beira Litoral podem ser analisados nas figuras F.14 a F.17.

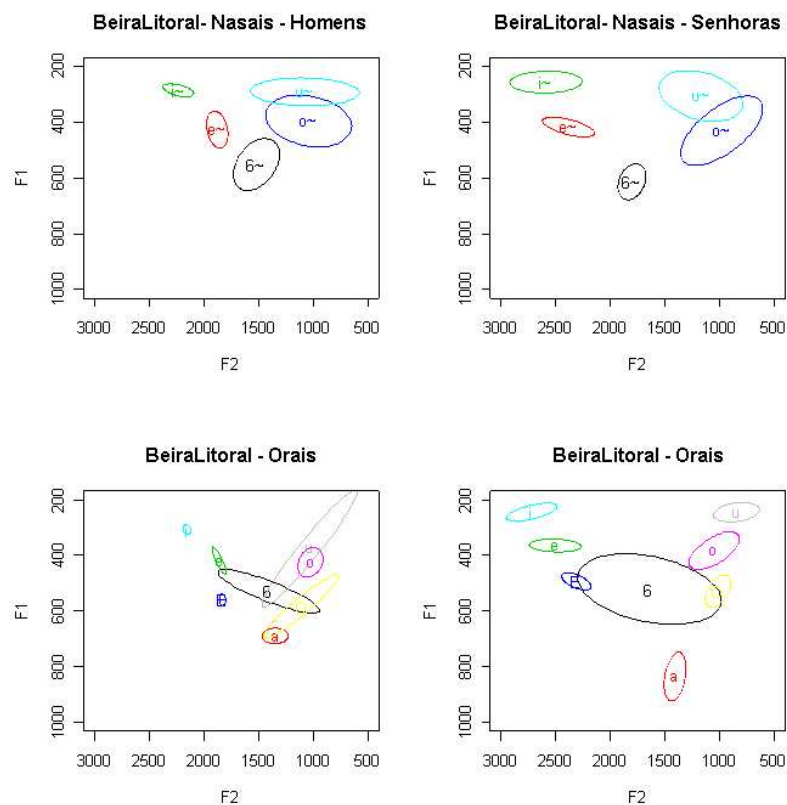


Figura F.14: Representação da distribuição das duas primeiras formantes para os falantes da BEIRA LITORAL.

As figuras F.18 a F.21 correspondem aos resultados obtidos para a região da Beira Interior.

Os resultados correspondentes às regiões do Alentejo e Algarve são apresentados, respectivamente nas figuras F.22 a F.25, e F.26 a F.29.

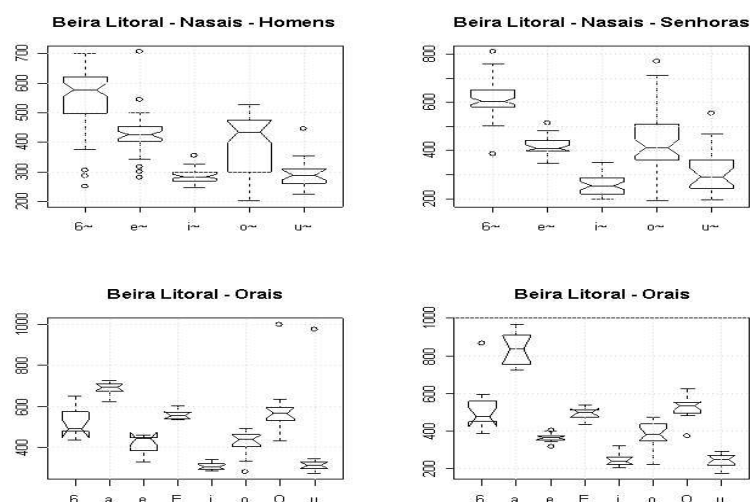


Figura F.15: Boxplots de F1 (separação por sexo e nasalidade) para os falantes da BEIRA LITORAL.

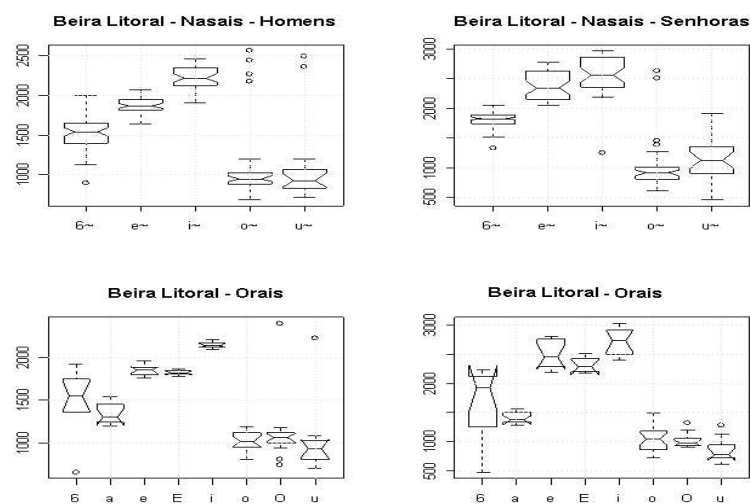


Figura F.16: Boxplots de F2 (separação por sexo e nasalidade) para os falantes da BEIRA LITORAL.

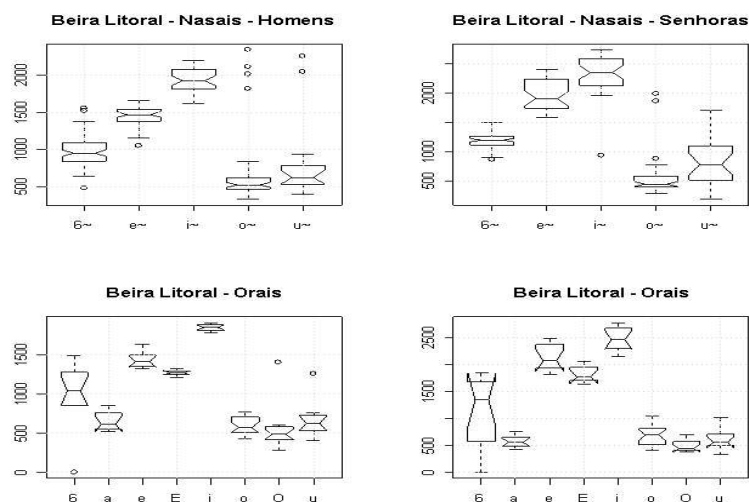


Figura F.17: Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes da BEIRA LITORAL.

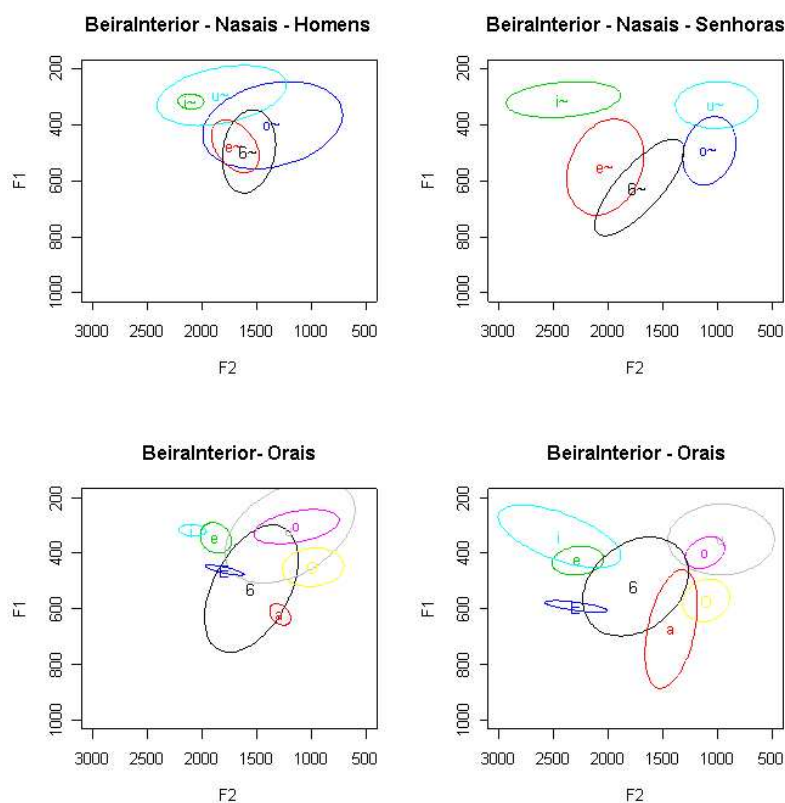


Figura F.18: Representação da distribuição das duas primeiras formantes para os falantes da BEIRA INTERIOR.

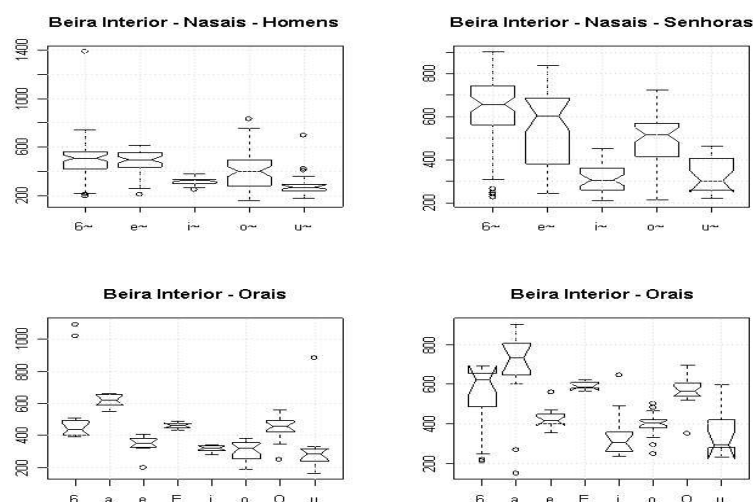


Figura F.19: Boxplots de F1 (separação por sexo e nasalidade) para os falantes da BEIRA INTERIOR.

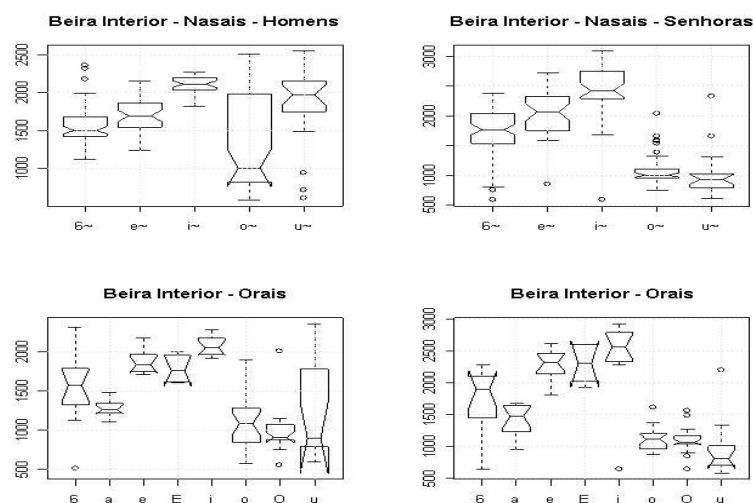


Figura F.20: Boxplots de F2 (separação por sexo e nasalidade) para os falantes da BEIRA INTERIOR.

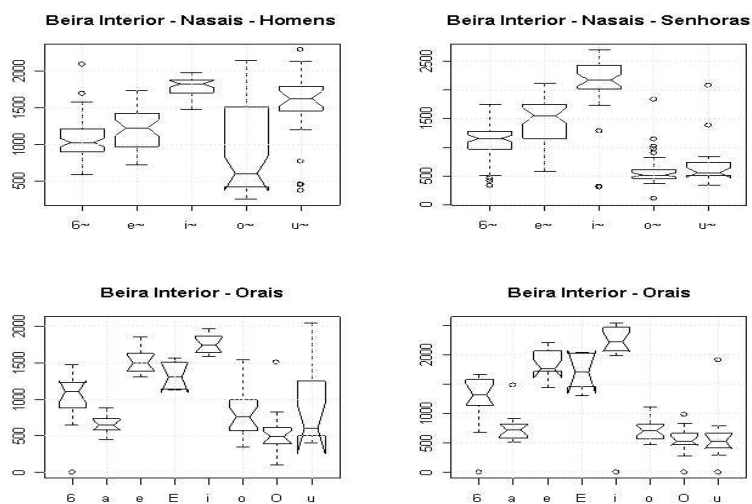


Figura F.21: Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes da BEIRA INTERIOR.

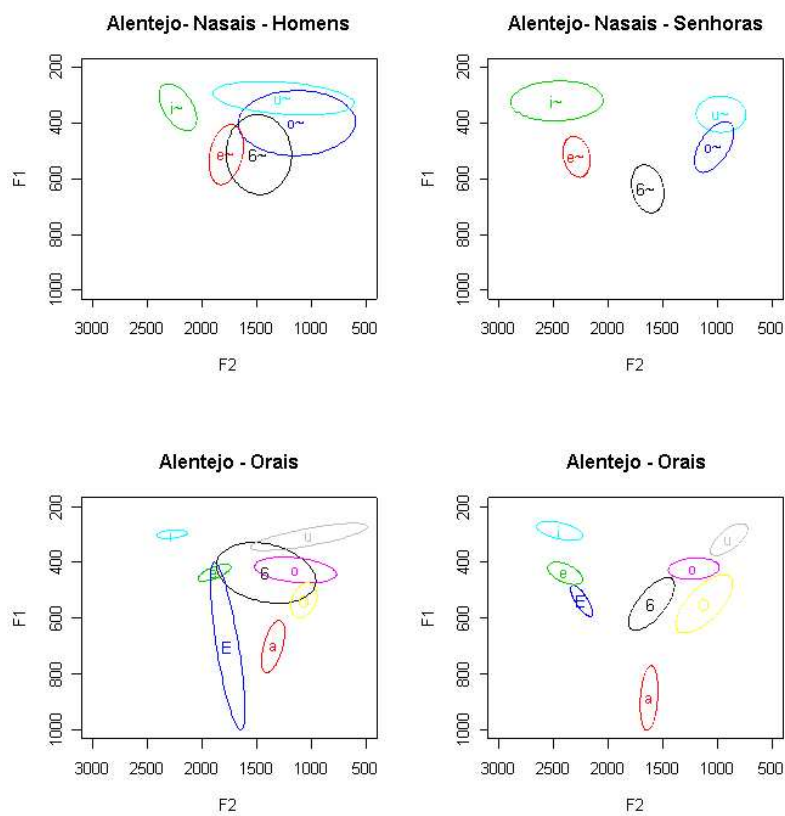


Figura F.22: Representação da distribuição das duas primeiras formantes para os falantes do ALENTEJO.

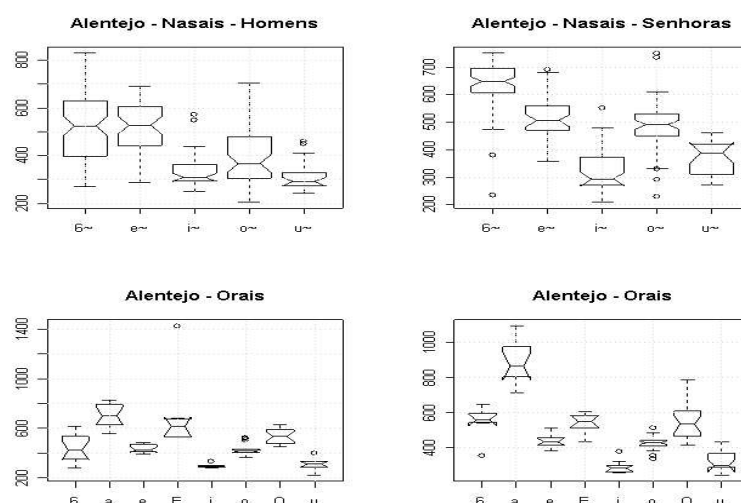


Figura F.23: Boxplots de F1(separação por sexo e nasalidade) para os falantes do ALENTEJO.

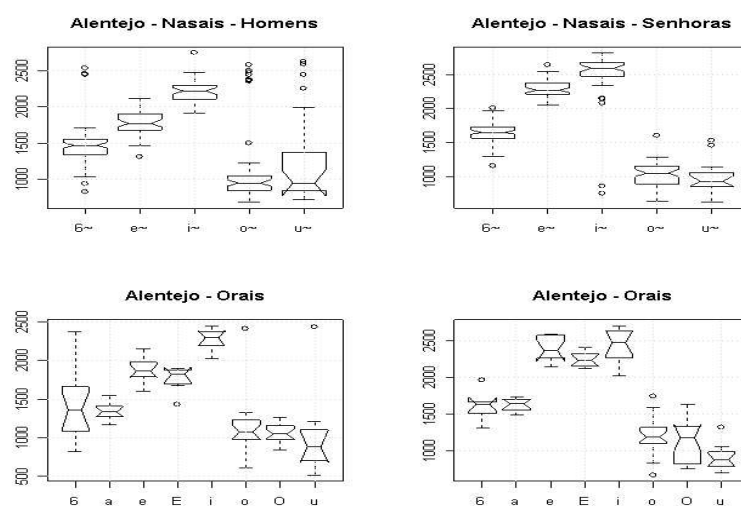


Figura F.24: Boxplots de F2 (separação por sexo e nasalidade) para os falantes do ALENTEJO.

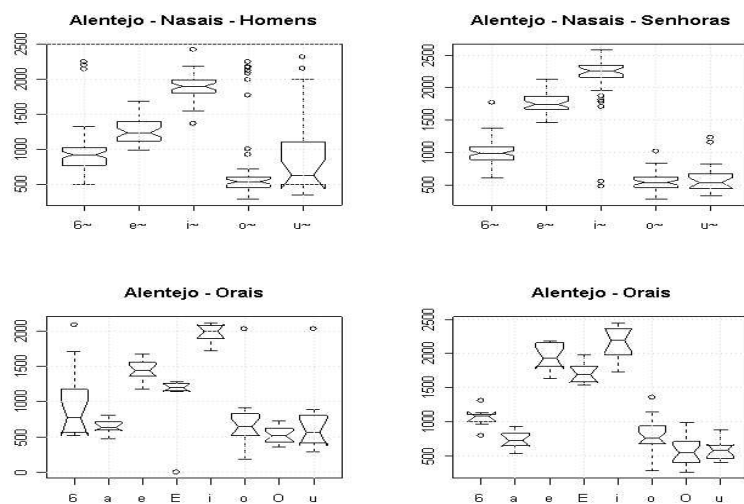


Figura F.25: Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes do ALENTEJO.

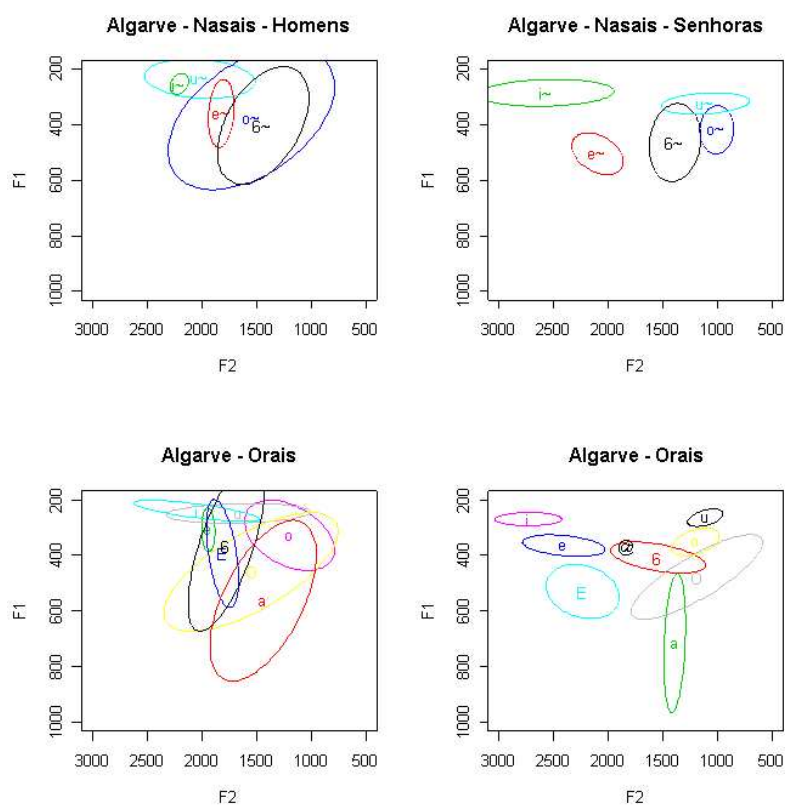


Figura F.26: Representação da distribuição das duas primeiras formantes para os falantes do MINHO.

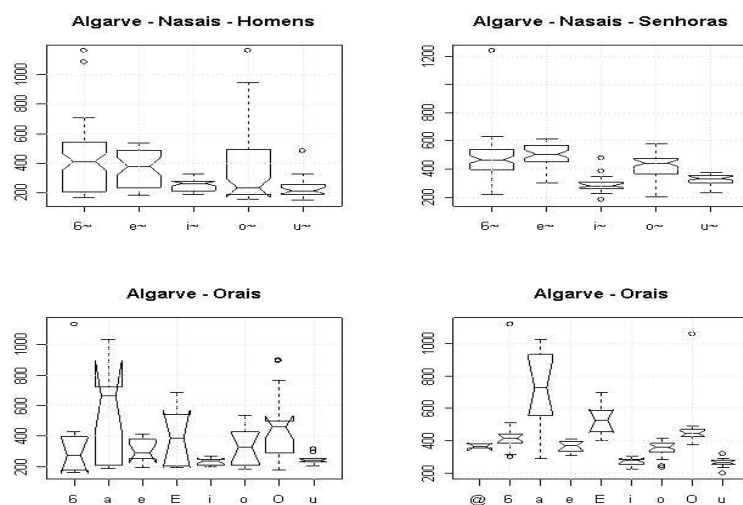


Figura F.27: Boxplots de F1 (separação por sexo e nasalidade) para os falantes do ALGARVE.

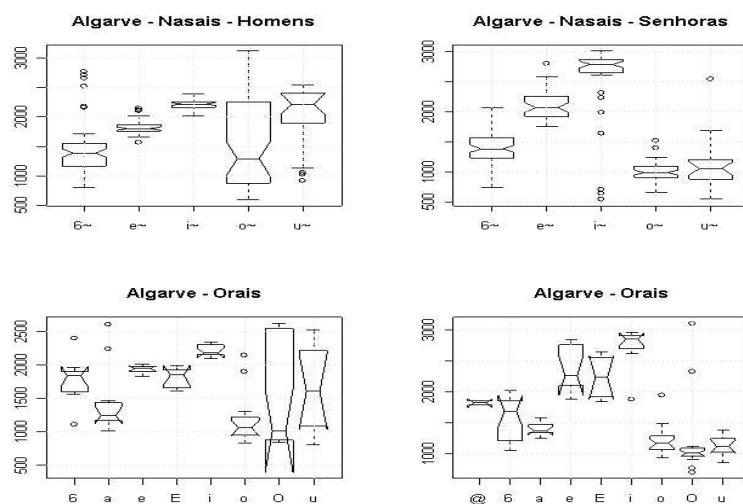


Figura F.28: Boxplots de F2 (separação por sexo e nasalidade) para os falantes do ALGARVE.

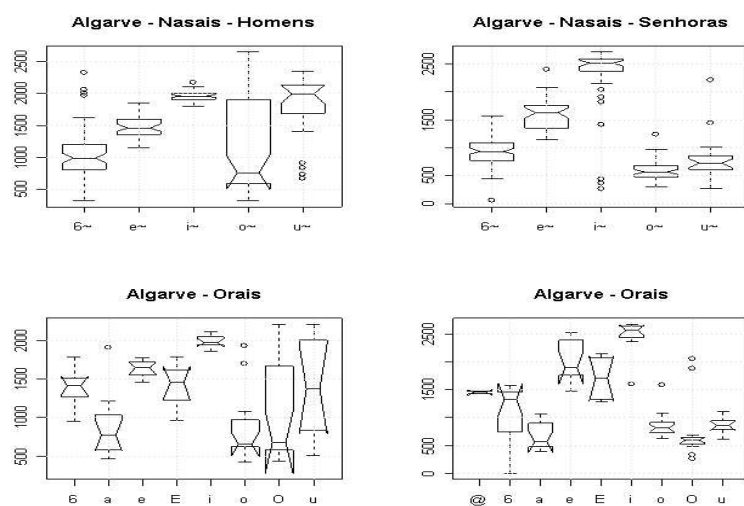


Figura F.29: Boxplots de F2-F1 (separação por sexo e nasalidade) para os falantes do ALGARVE.